

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 08:52:10 ; Search time 291 Seconds

(without alignments)
2925.772 Million cell updates/sec

Title: US-10-615-144-3

Perfect score: 1737

Sequence: 1 gaattcgccgcgcattga.....ataatcgccgcgaattc 1737

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4161359 seqs, 245077644 residues

Total number of hits satisfying chosen parameters: 8322718

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New:*

- 1: /cgn2_6/ptodata/1/pubna/US09_NEW_PUB.seq.*
- 2: /cgn2_6/ptodata/1/pubna/US06_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubna/US07_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubna/US08_NEW_PUB.seq.*
- 5: /cgn2_6/ptodata/1/pubna/PCT_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubna/US10_NEW_PUB.seq.*
- 7: /cgn2_6/ptodata/1/pubna/US11_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubna/US11_NEW_PUB.seq.*
- 9: /cgn2_6/ptodata/1/pubna/US11_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	146.4	8.4	3109	US-10-821-234-63	Sequence 63, Appl
C 2	41.6	2.4	977	US-10-750-185-58338	Sequence 58338, A
C 3	36.6	2.1	2725	US-10-131-826A-479	Sequence 479, App
C 4	35.4	2.0	1082144	US-11-117-187-211	Sequence 211, App
C 5	34.6	2.0	215308	US-11-121-086-77	Sequence 77, Appl
C 6	34.2	2.0	747	US-10-750-185-57746	Sequence 57746, A
C 7	34.2	2.0	80450	US-11-117-187-201	Sequence 201, App
C 8	34	2.0	1276	US-10-750-185-33227	Sequence 33227, A
C 9	34	2.0	3619	US-10-793-626-4157	Sequence 4157, Ap
C 10	34	2.0	645179	US-10-995-561-13293	Sequence 13293, A
C 11	33.8	1.9	1056	US-10-750-185-52662	Sequence 52662, A
C 12	33.8	1.9	2460	US-10-793-626-4439	Sequence 4439, Ap
C 13	33.8	1.9	2683	US-10-793-626-3709	Sequence 3709, Ap
C 14	33.8	1.9	3022	US-10-793-626-3969	Sequence 3969, Ap
C 15	33.8	1.9	4039	US-10-793-626-3361	Sequence 3361, Ap
C 16	33.6	1.9	794	US-10-750-185-52150	Sequence 52150, A
C 17	33.6	1.9	1432	US-10-750-185-56595	Sequence 56595, A
C 18	33.6	1.9	14341	US-10-995-561-13428	Sequence 13428, A
C 19	33.6	1.9	398287	US-10-995-561-13396	Sequence 13396, A
C 20	33.4	1.9	685	US-10-750-185-34681	Sequence 34681, A
C 21	33.4	1.9	1936	US-11-054-168B-15	Sequence 15, Appl
C 22	33.2	1.9	150437	US-11-112-908-44	Sequence 44, Appl
C 23	33.2	1.9	182314	US-11-112-908-45	Sequence 45, Appl

ALIGNMENTS

RESULT 1

US-10-821-234-63/C
; Sequence 63, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pt_seq_genes version 1.0
; SEQ ID NO 63
; LENGTH: 3109
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-821-234-63

Query Match	8.4%	Score 146.4;	DB 6;	Length 3109;
Best Local Similarity	51.1%	Pred. No. 7.3e-34;		
Matches 373;	Conservative	0;	Mismatches 351;	Indels 6;
Gaps 1;				

QY	676	TACTACAAAATTGCACGTCATTACAAAGTGGGCAATGATGACGTGTTTACAAAGATAAT	735
DB	1893	TACTACAAAGATCGCGCCACTACCGCTGGCGCTGGGCCAGGTCTTCGGCAGTTTCGC	1834
QY	736	TTTAGCGGTGTTATCATACTAGACATGATGAAATGCCCTGATTTTTTTCAGCTTT	795
DB	1833	TTCCCGCGCGCGCTGGTGGTGGAGATGACCTGGAGGTGGCCCTCGGACTTCTTCGAGTAC	1774
QY	796	TTTAGAGGTGGAGCTACTCTTCTTCACAGACAGACAGTCGATATGCTATTTCTTCTGG	855
DB	1773	TTTCGGGCAACCTATCCGCTGCTGAAGCCGACCCCTCCTGTGGTGGCTCTCGGCTGG	1714
QY	856	AATGACAATGGCAAAATGCAGTTTGTCCAAAG-----ATCCTTATGCTCTTTACCGCTCA	909
DB	1713	AATGACAAGCGCAAGAGCAGATGTTGGAGCGCCAGCAGCTGAGCTCTCTACCGCACC	1654
QY	910	GATTTTTTCCCGGCTTGGATGGATGCTTTCAAAATCTACTTGGGACCAATATCTCCA	969
DB	1653	GACTTTTCCCTGGCTGGCTGCTGTGTTGGCGAGCTCTGGGCTGAGCTGGAGCCC	1594

Db 1293 TGCTGTGGTCTGGAAGGACCTGGACATCTGCTGTGGATTTTTCAGTTTCTCTGAGCCA 1352
Qy 804 TGGAGCTACTCTTTCAGACAGACAAAGTCGATTCGCTATTCTCTTCCTGGAATGACAA 863
Db 1353 ATCCATCCACTACTGGAGGAGTACAGCCTGTACTGCATCTCTGCTGGAATGACCA 1412
Qy 864 TGGACAAATGAGTTTGTCAAGATCTTATGCTCTTTTACCGCTCAGATTTTTCCTCGG 923
Db 1413 GGGGTATGAACACACAGCTGAGGCCAGCAGCACTACTGACGCTGAGAGCCATGCTCTGG 1472
Qy 924 TCTTGGATGGATGCT 938
Db 1473 GCTGGCTGGGTGCT 1487

RESULT 4

US-11-117-187-211/c
; Sequence 211, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
; APPLICANT: PREUSS, DAPHNE
; APPLICANT: COPENHAVER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/11/117,187
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531,120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
; LENGTH: 1082144
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-11-117-187-211

Query Match 2.0%; Score 35.4; DB 7; Length 1082144;
Best Local Similarity 45.4%; Pred. No. 1e+02;
Matches 164; Conservative 0; Mismatches 196; Indels 1; Gaps 1;
Qy 498 TATTAAATCCATCTTAAATACCAATATCTGTTGCGTCAAAATATCTCTTTTCATATC 557
Db 497648 TAAATTATTAGGTGTAGAAATAACTAATAATTTTAAATATATTAATTTTGGATATTT 497589
Qy 558 CCAGATGGATCACAATCTGATGTGAGGAGCTTCTGCTTTGAGCTATGATCAGTCAGCTA 617
Db 497588 CGTTATCCAAATCTTTTGGTCTCTATCCATTCGGTTTCGGTTATTTTAAATTTATGTT 497529
Qy 618 TATGAGCACTGGATTTTGAACCTGTGCTACTGAAAGACAGCGGAGCTGATTT-GCAT 676
Db 497528 TCGGTTTCAGTTCGGATTCGGTTTAAAAATTTTAAAAACCGCATCGAATTCGATTCGGTT 497469
Qy 677 ACTACAAATTCACGCTCATACAGTGGCATTCGATCAGCTGTTTACAGACATAAT 736
Db 497468 ATTTTAAATTTATGTTTCGTTTCAGTTCGGATTCGGTTTAAATTAATATCTATTCCG 497409
Qy 737 TTAGCGGTGTTTATCATCTAGAAAGATGATATGAAATTTGCCCTGATTTTTCGACTTTT 796
Db 497408 TTATTTTGTTTAAATATTTTATGATATTTTAGTATTTTTCGATATTTTCATATATTTT 497349
Qy 797 TTGAGGCTGGAGCTACTCTTTTGAAGACAAAGTCGATATGCTATTTCTTCTTGA 856
Db 497348 GGTACTTTCGAGTATTTTACTATATACGTATATCTATTTTAAATTTAGTTTGA 497289
Qy 857 A 857
Db 497288 A 497288

RESULT 5

US-11-121-086-77/c

; Sequence 77, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 77
; LENGTH: 215308
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-77

Query Match 2.0%; Score 34.6; DB 7; Length 215308;
Best Local Similarity 52.4%; Pred. No. 75;
Matches 76; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
Qy 1017 TCACAGAGTCGACAAATTTATTCGCCAGAGTTTCCGCCAGACATATATTTTGGTGACCA 1076
Db 91543 TTACAAATTTTGTCCCTTCTTCTGTTTCTAAATTTTGTAAATATACCATTTTCTCTGATT 91484
Qy 1077 TGGTCTAGTTGGGCGAGTTTTCAGACGATATCTTGAGCAATTAACCTAAATGATGT 1136
Db 91483 CCTTAGTGTTCCTCCCATGTTTCTTTTACTCTTTGAGCAAAATTTAAGACAATTTGTTT 91424
Qy 1137 CCAGTTGATGGAAGTCAATGGAC 1161
Db 91423 TAAAGTCTTTGTAAGTCAATGTC 91399

RESULT 6

US-10-750-185-57746/c
; Sequence 57746, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57746
; LENGTH: 747
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-57746

Query Match 2.0%; Score 34.2; DB 6; Length 747;
Best Local Similarity 55.5%; Pred. No. 3.1;
Matches 66; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
Qy 1275 TGTGCTATTTCAGTACAGAGATCACTAGACTTTGAAAATATCGCACCGCAATTTGGCAT 1334
Db 368 TTTATGTTTGAACACAGACACTAAATGGACATTTTAAATGTCGATCGACAGTTTGTCTCT 309
Qy 1335 TTTTGAAGAATGGAAGGATGTTTACCAGCTGACGATATAAAGGAATAGTAGTTTTC 1393
Db 308 TTTATTTGTGAATATGGATAATCTTGTCTTTCAAAGGGTTTAGGTAGACTGGCTTTC 250

```
RESULT 7
US-11-117-187-201
; Sequence 201, Application US/11117187
; Publication No. US20050266560A1
; GENERAL INFORMATION:
; APPLICANT: PREUSS, DAPHNE
; APPLICANT: COPENHAVER, GREGORY
; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS
; FILE REFERENCE: ARCD:309US
; CURRENT APPLICATION NUMBER: US/11/117,187
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: US/09/531,120
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,219
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 201
; LENGTH: 80450
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-11-117-187-201

Query Match      2.0%; Score 34.2; DB 7; Length 80450;
Best Local Similarity 54.3%; Pred. No. 55;
Matches 69; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 1087 TTGGGCGAGTTTTCACAGCAGTATCTTGAGCCAAATTAACCTAAATGATGTCCAGGTGAT 1146
Db 13727 TAGGTTAATTATTCACCGACGAATTAATCTTAGAGAGACAAATGAAGGTTCAAGTGGAG 13786

Qy 1147 TGAAGTCAATGGACCTTAGTTACCTTTTGAGGACCAATTAACGTGAAACACTTTGGTGAC 1206
Db 13787 TCAAAGACCAAGGACCGAGTCACCTTATTATGGGAATATGAGTGTGGCACTTTTTCCTCC 13846

Qy 1207 TTGGTTA 1213
Db 13847 TTATTTA 13853

RESULT 8
US-10-750-185-33227
; Sequence 33227, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33227
; LENGTH: 1276
; TYPE: DNA
; ORGANISM: Bovine 19866880583544
US-10-750-185-33227

Query Match      2.0%; Score 34; DB 6; Length 1276;
Best Local Similarity 52.1%; Pred. No. 4.9;
Matches 76; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 942 AAATCTACTTGGGACGAATTTATCTCAAAGTGCCAAAGGCTTACTGGGACGACTGGCT 1001
```

```
Db 1088 ATAAGCTCATCTAGACTTAACTTCTCCAAAGTGCCACAGTAATTCAGATTTCTTTGCT 1147
Qy 1002 AGACTCAAAGAGAATCACAGAGTTCGACAAATTTATTCGCCCGCAGAGTTTTCGACACATA 1061
Db 1148 AAGATTAAACACACTTGTTCAGACAAACACATATTTCTGTACCTATAAATTTTGTGCATACA 1207
Qy 1062 TAATTTTGGTGAGCATGGTTCTTAGTT 1087
Db 1208 TTGCGTGAGTGTCTCCCGATGCAAAAT 1233

RESULT 9
US-10-793-626-4157/c
; Sequence 4157, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4157
; LENGTH: 3619
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4157

Query Match      2.0%; Score 34; DB 6; Length 3619;
Best Local Similarity 50.0%; Pred. No. 9.4;
Matches 85; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Qy 1213 AAAAGGCTAAGCCCATCCATGGAGCTGATGCTGCTTGAAGCATTTAAACATAGATGGT 1272
Db 2960 ACAATCCTGCACCTAGACTCAGAAAGGTAATTCGTATGGTAACTATTATATGGCAAGAC 2901

Qy 1273 GATGTCGTATTTCAGTACAGAGATCAACTAGACTTTGAAATATCCACGGCAATTTGGC 1332
Db 2900 GATATGACGATTGATGCTGATTTTATCAACACACCTAAATACAGCTTAGGGAACATATGTA 2841

Qy 1333 ATTTTGAAGAAATGGAAGGATGGTGATACCAAGTGTGATGATGATGATGATGATGATGAT 1382
Db 2840 TGGTATGACACTAATAAAGATGGTATTCAAGGTGATGATGATGATGATGATGATGATGAT 2791

RESULT 10
US-10-995-561-13293
; Sequence 13293, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13293
; LENGTH: 645179
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-13293

Query Match      2.0%; Score 34; DB 6; Length 645179;
```



```

Best Local Similarity 46.5%; Pred. No. 2e+02;
Matches 106; Conservative 1; Mismatches 121; Indels 0; Gaps 0;
QY 1165 AGTACCTTTGGAGGACAAATACGTGAACACATTTGGTGACTGGTTAAAGAGCCTAAG 1224
DB 149618 ATTTACATGGAAGAAGTCTACTGTGTACCAAGTTGCATATTTTGTGTAAAGTTCTCAG 149677
QY 1225 CCCATCCATGAGCTGATGCTCTTGAAGCATTTTACATAGATGGTGATGCTGATTT 1284
DB 149678 CCATGGTATGATGATCAATTTGGTGTGTGACGCCATCTCTCAGACCTTTATTTCTCTGTT 149737
QY 1285 CAGTACAGAGATCAACTAGACTTTGAAATATCGCAGCGCAATTTGGCAATTTTGAAGAA 1344
DB 149738 CAGCCATACATAAATGACATTTTCAATGTTCCAGGTAGTTTCAATTTTAAAGATAA 149797
QY 1345 TGAAGAGATGGTGTACCAAGTGCAGCATATTAAGGAATAGTAGTTTTC 1392
DB 149798 CAGAAAACCTAGTATTCAATGGTTAAATAAAGCAATACCTATC 149845

```

RESULT 11

```

US-10-750-185-52662
; Sequence 52662, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52662
; LENGTH: 1056
; TYPE: DNA
; ORGANISM: Bovine 1986688139749
US-10-750-185-52662

```

```

Query Match
Best Local Similarity 1.9%; Score 33.8; DB 6; Length 1056;
Matches 62; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 142 TTTTGTGTGATTCGGTACCTCTCATCTTGGCTGTGCTGCTTCACTACACAG 201
DB 370 TTTGGCAGACATTTCCCTCTCCCTCTCAATCTCTCTCCCTCACCACCCATCCAGACGAG 429
QY 202 ATCCGCTTTTTCGACAGTGCAGATATGAGATGCGCTTGTGCTG 250
DB 430 TTAAAGGATTTGCTTTCGAGAGAGAAATGAACCGCTCTTTGCTTCTG 478

```

RESULT 12

```

US-10-793-626-4439/c
; Sequence 4439, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 4439
; LENGTH: 2460
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4439

```

```

Query Match
Best Local Similarity 1.9%; Score 33.8; DB 6; Length 2460;
Matches 170; Conservative 0; Mismatches 207; Indels 2; Gaps 1;
QY 884 AAGATCCCTTATGCTCTTTACCGCTCAGATTTTTCGCCGGTCTTGGATGGATGCTTTCAA 943
DB 1189 AAAATTTATTAACCAATAACATAGATACATTTTTCATTAAGTTTCACTAGTTTTCAAA 1130
QY 944 AATCTACTTGGGACGAATATCTCAAGTGGCAAGGCTTACTGGGACGACTGGCTAA 1003
DB 1129 AAAATTTTATATTTTATGATACTTAAAGTGATATATTTTCTACATAATATTTAA 1070
QY 1004 GACTCAAGAGCAATCACAGAGGTGACAAATTTATTTGCCCCAGAACTTTGCAGAACATATA 1063
DB 1069 TTTTCAAGGAGATGTAAAGTTGAAAAATTTGCAAAAATTAATTTTGGTGGCATTTT 1010
QY 1064 ATTTTGGTGAAGCATGGTTCTTAGTTTGGGGCAGTTTTCACAGCAGTATCTTGAGCCAAATA 1123
DB 1009 AGTATCGGGTCCAGGATAGCGAGTGTACAAACAAATATACTACGCCAAGAAAGTCA 950
QY 1124 --AATAAATGATGTCAGGTTGATGGAAGTCAATGACCTTAGTTACCTTTTGGAGGA 1181
DB 949 CGATTCAACTCTCAAAATATTAATTAAGTGGGAACGTATGATATCTTCTCAAGTTGATTC 890
QY 1182 CAATTACGTGAACACTTTTGGTGAAGTCTTAAAGGCTAAGCCCATCCATGGAGCTGA 1241
DB 889 CAAACGATGAACAAATTTAAAGAAATAAGAAAGAAAGATAATAATTTCCACATAACTAA 830
QY 1242 TGCTGCTCTTGAAAGCATTT 1260
DB 829 ACATGGAATAAAGTCGTT 811

```

RESULT 13

```

US-10-793-626-3709
; Sequence 3709, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3709
; LENGTH: 2683
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3709

```

```

Query Match
Best Local Similarity 1.9%; Score 33.8; DB 6; Length 2683;
Matches 170; Conservative 0; Mismatches 207; Indels 2; Gaps 1;
QY 884 AAGATCCCTTATGCTCTTTACCGCTCAGATTTTTCGCCGGTCTTGGATGGATGCTTTCAA 943
DB 37 AAAATTTATTAACCAATAACATAGATACATTTTTCATTAAGTTTCACTAGTTTTCAAA 96

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 08:02:40 ; Search time 6397 Seconds
(without alignments)
12704.262 Million cell updates/sec

Title: US-10-615-144-3
Perfect score: 1737
Sequence: 1 gaattcgccgcgcattga.....ataatcgccgcgcgaattc 1737

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	739.6	42.6	894	8	DN981229 SV6_35B12
2	723.4	41.6	925	5	BQ505949 EST613364
3	722.4	41.6	1654	4	CNS0A4J5 Arabidops
4	630	36.3	768	2	BG594788 EST493466
5	599.6	34.5	772	5	BQ505948 EST613363
6	593	34.1	724	2	B1179808 EST520753
7	580.2	33.4	642	2	BG592816 EST491494
8	578.6	33.3	643	1	AI486771 EST245093
9	553.6	31.9	661	2	BG889872 EST515723
10	537.8	31.0	673	1	AW979500 EST310521
11	527	30.3	659	1	AW735827 EST336595
12	517.2	29.8	843	3	BM779745 EST590321
13	514	29.6	574	1	AW931180 EST357023
14	510.6	29.4	758	1	AJ806862 AJ806862
15	508.2	29.3	852	3	BM780056 EST590632
16	496.6	28.6	874	8	DR927057 EST111859
17	484.8	27.9	1036	5	BX839862 EX839862
18	465.6	26.8	875	5	DR015231 STRS1_2_F
19	456.8	26.3	705	7	CO491834 G.h.fbr-s
20	453.8	26.1	837	8	DR692299 EST108238
21	445.6	25.7	674	5	BUB37218 T96809 P
22	444.4	25.6	454	3	BP527408 BP527408

ALIGNMENTS

RESULT 1
DN981229
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

DN981229 894 bp mRNA linear EST 15-JUN-2005
SV6_35B12_SV6_35t_093 SV6 Solanum chacoense cDNA, mRNA sequence.
DN981229 GI:67769040
EST.
Solanum chacoense (Chaco potato)
Solanum chacoense
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Solanum.
1 (Bases 1 to 894)
Germain,H., Rudd,S., Zotti,C., Caron,S., O'Brien,M., Chantha,S.-C.,
Lagace,M., Major,F. and Matton,D.P.
A 6374 unigene set corresponding to low abundance transcripts
expressed following fertilization in Solanum chacoense Bitt., and
expression profile of 30 receptor-like kinases
Unpublished (2005)
Contact: Daniel P. Matton
Daniel P. Matton
Universite de Montreal
Institut de recherche en biologie vegetale, 4101 rue Sherbrooke
est, Montreal, QC, Canada, H1X 2B2
Tel: 5148723967
Fax: 5148729406
Email: dp.matton@umontreal.ca.
Location/Qualifiers
1.894
/organism="Solanum chacoense"
/mol_type="mRNA"
/db_xref="taxon:4108"
/tissue_type="2-6 days post-pollination ovaries"
/dev_stage="Zygotic to globular embryo"
/lab_host="SOLR"
/clone_lib="SV6"
/note="Organ: Deparicarped ovary/ovules with placenta
tissue; Vector: pBS SK; Site 1: EcoRI; Site 2: XhoI;
Deparicarped ovaries were isolated from a compatible cross
between self-incompatible accessions having the S12S14
self-incompatibility alleles (G4 line as female) and
S11S13 self-incompatibility alleles (V22 line as male).
Those two lines were derived from the parental lines PI
458314 (which carries the S11 and S12 self-incompatibility
alleles) and line PI 230582 (which carries the S13 and

FEATURES
source

Qy	738	TAGCCGTGTTATCATCTACGAGATGATATGGAATTCGCCCTGATTTTTTTTGACTTTTT	797
Db	624	TAGTCGAGTGTATATCTACGAGAGATGATATGGAATTCGCCGACATCTCTTGATTACTT	683
Qy	798	TGAGGCTGGAGCTACTCTCTTTCACACAGACACAGTCGATATATGCTATTTCTCTCTCGAA	857
Db	684	TGAGGCTGCGAGCTAGTCTCATGGATAGGGATAAAACCATTATGGCTGCTTCATCATGGAA	743
Qy	858	TGACAAATGGACAAATGCAGTTTGTCCAAAGATTCCTTATGCTCTTTACCGCTCAGATTTTTT	917
Db	744	TGATAATGGACAGAAGCAGTTTGTGCATGATCCCTATGCGCTATACCGATCAGATTTTTT	803
Qy	918	TCCCGTCTTGGATGGATGCTTTCAAATCTACTTTGGAGCAATATCTCCAAAGTGGCC	977
Db	804	TCCTGGCCTTGGGTGGATGCTCAAGAGATGCTCTTGGGATGAGTTATCACCAAGTGGCC	863
Qy	978	AAAGGCTTACTGGACGACACTGGCTAAGACCTCAAAGAGAACTACAGAGGTCGACAAATTTAT	1037
Db	864	AAAGGCTTACTGGGATGATTGGCTTGAAGCTAAAGGAAACCATAAAGCCGCCCAATTCAT	923
Qy	1038	TCGCCCAAGAGTTTGCAGAACATATAATTTTGGTGAGCATGGTTCTAGTTTGGGCGAGTT	1097
Db	924	TCGACCGGAGTCTGTAGAACATACAAATTTTGGTGAACATGGGTCTAGTTTGGTCAATT	983
Qy	1098	TTTCAAGCAGTATCTTGAGGCAATTAACATAAATGATGTCAGGTTGATTTGGAAGTCAAT	1157
Db	984	TTTCAGTCAGTATCTGATCCTATAAGTTTATACGATGACGGTTGATTTGGAATCAAA	1043
Qy	1158	GGACCTTAGTTACCTTTTGGAGGACAATTAGCTGAACACATTTTGGTGACTTGGTTAAAA	1217
Db	1044	GGACCTGGGATCTTGACAGAGGTTACTTTTACCAAGTACTTTTCTGGCTTAGTGAGATA	1103
Qy	1218	GGCTTAAGCCCATCCATGGAGCTGATGCTGCTCTTCAAGACATTTTAAACATAGATGGTATGT	1277
Db	1104	AGCAGCTAATTCAGGTTCTGACCTTGCTTAAAGGCTCAAAACATAAAGTATGATGT	1163
Qy	1278	CGGTATTTCACTACAGACATCAACTAGACTTTTGAATAATCGCAGGCAATTTGGCATTTTT	1337
Db	1164	TCATTATCCGGTATAAAGCAAGTAGAGTTTGAACGCATTGACGGGAAATTTTGTATATT	1223
Qy	1338	TGAAGAATGGAAGATGGTGTACCACTGCAGGCATATAAAGGAATAGTAGTTTCCGGTA	1397
Db	1224	TGAAGATTTGAAGGATGGTGTGCTCTTAACAGACATATAATTGGAGTAGTGGTGTTTGAAAT	1283
Qy	1398	CCAAACGTCCAGACGTGATTCTCTTGTGTGGCCATGATTTCGCTTCAACAACTCGGAATTTGA	1457
Db	1284	CCAGATAACAGACGTGATTCTCTTGGTTGGGCCAGATCTGTATTGCTGCTTGGAAATCG	1343
Qy	1458	AGATATTAAACAAAGATATGATTCGACGAGCGCCGGCAAAATTTTTGACTTA	1509
Db	1344	AAATTCCTGTATGCAAAACATGACTCTCTCCAAACGAGCAATATTTTCATCTTA	1395

RESULT 4	
BG594788	
LOCUS	768 bp mRNA linear EST 07-MAR-2003
DEFINITION	EST493466 cSTS Solanum tuberosum cDNA clone cSTS8020 5' sequence, mRNA sequence.

```

FEATURES
    source
        1..768
            /organism="Solanum tuberosum"
            /mol_type="mRNA"
            /cultivar="Kennebec"
            /db_xref="taxon:4113"
            /clone="cSTS8020"
            /tissue_type="sprouting eyes from tubers"
            /dev_stage="12-14 weeks post harvest"
            /lab_host="SOLR"
            /clone_lib="cSTS"
            /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
            XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
            taken from tubers. The tubers were incubated at 26C in the
            dark for 2-3 weeks prior to sprouting. The eyes were
            frozen in liquid nitrogen immediately upon removal from
            tubers."

```

COMMITTEE


```

QY 786 TTTTGACTTTTGGAGCTGGAGCTACTCTCTT 819
DB 733 TTTTGACTATTGGAGGCTGAGCTACTCTCTT 766

RESULT 5
BQ505948 772 bp mRNA linear EST 07-MAR-2003
LOCUS EST613363 Generation of a set of potato cDNA clones for microarray
DEFINITION analyses mixed potato tissues Solanum tuberosum cDNA clone STMGH33
5' end, mRNA sequence.
ACCESSION BQ505948
VERSION BQ505948.2 GI:21921855
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 772)
AUTHORS Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
Karamycheva,S.A.
Generation of a set of potato cDNA clones for microarray analyses
Unpublished (2002)
On Jun 10, 2002 this sequence version replaced gi:21364817.
Other ESTs: EST613364
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3.
Location/Qualifiers
1..772
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec or Bintje"
/db_xref="taxon:4113"
/clone="STMGH33"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/clone_lib="Generation of a set of potato cDNA clones for
microarray analyses mixed potato tissues"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Combination of untreated and Phytophthora
infestans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating
eyes, tubers, or roots."

ORIGIN
Query Match 34.5%; Score 599.6; DB 5; Length 772;
Best Local Similarity 93.9%; Pred. No. 9,3e-160;
Matches 657; Conservative 0; Mismatches 39; Indels 4; Gaps 3;

QY 126 GATGAGAGGAAACAAGTTTGTGTGATTCGGTACCTTCCTCACTTGGCTGCTGCTCGC 185
DB 73 GATGAGAGGAAACAAGTTTGTGTGATTCGGTACCTTCCTCGTGGCTGCTGCTCGC 132

QY 186 CTTTCATCTACACAGATCGGCTTTTGGACAGATCAGATATGCAATGCGCTTGC 245
DB 133 CTTTCATCTACAGATCGGCTTTTGGACAGATCAGATATGTAAGCGCTTGC 192

QY 246 TGTGCAATTGAAGCAGAAAATCATTTGTAACAGCCAGACAGATCTCTTATTGACAGAT 305
DB 193 TGTGCAATTGAAGCAGAAAATCATTTGTAACAGTGAACAGATCTCTTATTGACAGAT 252

QY 306 TAGCCTTCAGCAGGGAAGATAGTTGCTCTTTGAGAACAAATGAAGCGTCAGGACAGGA 365
DB 253 TAGCCAGCAGCAGGGAAGAGTAGTAGTCTCTTGAAGAACAAATGAAGCGTCAGGACAGGA 312

366 GTGCCGCAAAATTAAAGGCTCTTGTTCAGGATCTTTGAAAGTAAGGGCATAAAAAGTTGAT 425
313 GTGCCGCAAAATTAAAGGCTCTTGTTCAGGATCTTTGAAAGTAAGGGCATAAAAAGTTAAT 372
426 CGGAAATCTACAGATGCCAGTGGCTGTAGTTGTTATGCTTGCATCGCAATCGGCTGATTA 485
373 CGGAGATGTGCAGATGCCAGTGGCAGCTGTAGTTGTTATGCTTGCAGTGTGCTGACTA 432
486 CCTGGAAGAAAGACTATTAAATCCATCTTTAAATACCAATATCTCTTGGCTCAAAATATCC 545
433 CCTGAGAGGAGACTATTAAATCCATCTTTAAATACCAATATCTCTTGCATCAAAATATCC 492
546 TCTTTTCATATCCAGGATGATCACATCTCTGATGTCAGGAGCTTCTTTGAGCTATGA 605
493 TCTTTTCATATCCAGGATGATCAAAATCCCTGATGTAAGAAAGCTTCTTTGAGCTATGA 552
606 TCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGTCATCTGAAAGACCA-GGGG 664
553 TCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGTCATCTGAAAGACCAAGGGG 612
665 AGCTGATTGCATCTACAAAATTCGACGTCAATTCACAAAGTGGG--CATTTGATCAGCTGTT 722
613 AACTGTTGTCATCTACAAAGATTCACGTCATTACAAAGTGGGCACTTTGGATCAGCTGTT 672
723 TTACAAGCATAA-TTTTAGCGCTGTTATCATCTACTAGTAGATGATATGAAAATTGGCCCTG 781
673 TCACAAGCATAAATTTTAGCGCTGTTATCATCTACTAGTAGATGATATGAAAATTGGCTGCTG 732
782 ATTTTGTGACTTTTGTGAGGCTGGAGCTACTCTCTCTTGA 821
733 ATTTTGTGACTATTTTGAGGCTGGAGCTACTCTCTCTTGA 772

RESULT 6
B1179808 724 bp mRNA linear EST 10-MAR-2003
LOCUS EST520753 cSTE Solanum tuberosum cDNA clone cSTE20K20 5' sequence,
DEFINITION mRNA sequence.
ACCESSION B1179808
VERSION B1179808.1 GI:14645619
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 724)
AUTHORS van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R., Cho,J.,
Chiemingo,A., Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and
Baker,B.
Generation of ESTs from in vitro grown microtubers
Unpublished (2001)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13F-R.
Location/Qualifiers
1..724
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Bintje"
/db_xref="taxon:4113"
/clone="cSTE20K20"
/tissue_type="axillary buds of stem explants; growing
sink-tubers"
/dev_stage="7, 8 and 10 days"
/lab_host="SOLR"
/clone_lib="cSTE"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:

```


XhoI; Tissue supplied by Christian Bachem and Richard Visser (Department of Plant Breeding, Wageningen University, The Netherlands). The cSTA libraries will attempt to capture the induction and initiation/initial growth of the tuber in an in vitro system as described in Bachem et al. (Plant Journal, 1996). Small microtubers develop from axillary buds attached to stem explants when placed on a high sucrose medium (10%). Visible morphological changes occur synchronously at day five in the axillary buds. The first library, cSTA (1-20) consists of axillary buds harvested on days 1-3. This targets those genes involved in induction of the microtubers. The following libraries, cSTA (21-40) and cSTA (41-60), capture genes involved in tuber initiation and outgrowth. This library is noted as P3 in Tanksley lab notebooks."

ORIGIN

Query Match 34.1%; Score 593; DB 2; Length 724;
Best Local Similarity 93.9%; Pred. No. 7.1e-158; Indels 0; Gaps 0;
Matches 617; Conservative 0; Mismatches 40;

QY 126 GATGAGGGGAACAGTTTTCGTGTGATTTCCGGTACCTCTCATCTTGGTGCTGCGC 185
DB |||||
QY 68 GATGAGGGGAACAGTTTTCGTGTGATTTCCGGTACCTCTCTCGTGGCTGCTCGC 127
DB |||||
QY 186 CTTTCATCTACACAGATCGCGCTTTTTCGCACACAGTCAGAAATGACAGATCGCCTTGC 245
DB |||||
QY 128 CTTTCATCTACACAGATCGCGCTTTTTCGCACACAGTCAGAAATGATGACCGCCTTGC 187
DB |||||
QY 246 TGCTGCAATTGAAGCAGAAATCATTTGACAGCCAGCAGATGCTTTATTGACACAGAT 305
DB |||||
QY 188 TGCTGCAATTGAAGCAGAAATCATTTGACAGTCAGACAGATGCTTTATTGACAAAGAT 247
DB |||||
QY 306 TAGCCTGCAGCAAGGAATAGTTGCTCTTTGAAGAACAAATGAAGCGTCAGGACACAGGA 365
DB |||||
QY 248 TAGCAGCAGCAGGAAGAGTAGTAGCTCTTGAAGAACAAATGAAGCATCAGGACACAGGA 307
DB |||||
QY 366 GTGCGGCAATTAAGGGCTCTTTTCAGGATCTTGAAGTAAGGCATATAAAAGTTGAT 425
DB |||||
QY 308 GTGCGGCAATTAAGGGCTCTTTTCAGGATCTTGAAGTAAGGCATATAAAAGTTGAT 367
DB |||||
QY 426 CGGAAATGTCAGATGCCAGTGGCTGTGATGTTGTTATGGCTTGCATCGGCTGATTA 485
DB |||||
QY 368 CGGAGATGTGAGATGCCAGTGGCAGTGTAGTTGTTATGGCTTGCATCGCTACTGACTA 427
DB |||||
QY 486 CTTGAAAGACTATTAATCCATCTTAAATATCAAAATATCTGTGCGTCAAAATATCC 545
DB |||||
QY 428 CTTGAGAGGACTATTAATCCATCTTAAATATCAAAATATCTGTGATCAAAATATCC 487
DB |||||
QY 546 TCTTTTTCATATCCAGGATGGATCACATCCCTGTGATGTCAAGAGCTTGTGAGCTATGA 605
DB |||||
QY 488 TCTTTTTCATATCCAGGATGGATCAATCCCTGTGATGTAAAGAGCTTGTGAGCTATGG 547
DB |||||
QY 606 TCAGCTGACGTATATGACAGCTTGGATTTTGAACCTGTGCTATCTGAAGACACAGGGA 665
DB |||||
QY 548 TCAGCTGACGTATATGACAGCTTGGATTTTGAACCTGTGCTATCTGAAGACACAGGGA 607
DB |||||
QY 666 GCTGATTGCATCTACAAAATTTGACAGCTCATTAACAAGTGGCATTTGGATCAGCTGTTTTA 725
DB |||||
QY 608 ACTGGTGCATCTACAGATTTGACGCTCATTAACAAGTGGCATTTGGATCAGCTGTTTCA 667
DB |||||
QY 726 CAAGCATAAATTTAGCCGTGTTATCATACTAGAGATGATATGGAATTTGCCCTGTA 782
DB |||||
QY 668 CAAGCATAAATTTAGCCGTGTTATCATACTAGAGATGATATGGAATTTGCTGCTGA 724
DB |||||

RESULT 7
BG592816
LOCUS BG592816 642 bp mRNA linear EST 07-MAR-2003
DEFINITION EST491494 cSTS Solanum tuberosum cDNA clone cSTS29 5' sequence,
mRNA sequence.
ACCESSION BG592816
VERSION BG592816.1 GI:13610956

KEYWORDS
SOURCE

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Solanum tuberosum (potato)
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 642)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chieningo,A.,
Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
Generations of ESTs from sprouting potato eyes
Unpublished (2000)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13F-R.

FEATURES
source

1..642
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS2B9"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/clone_lib="cSTS"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."

ORIGIN

Query Match 33.4%; Score 580.2; DB 2; Length 642;
Best Local Similarity 94.1%; Pred. No. 3.1e-154;
Matches 603; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 420 GTTGATCGGAAATGTACAGATGCCAGTGGCTGTAGTTGTTATGGCTTGCATCGGC 479
DB |||||
QY 2 GTTAAATCGGAGATGTGCAGATGCCAGTGGCAGCTGTAGTTGTTATGGCTTGCATCGTC 61
DB |||||
QY 480 TGATTACCTCGGAAAGACTATTAAATCCATCTTAAATACCAATATCTGTGCGTCAAA 539
DB |||||
QY 62 TGACTACCTGGAGAGGACTATTAAATCCATCTTAAATACCAACATCTGTTCATCAA 121
DB |||||
QY 540 ATATCCTCTTTTCATATCCAGGATGGATCAATCCTGTGATGTGAGAAAGCTTGTGAG 599
DB |||||
QY 122 ATATCCTCTTTTCATATCCAGGATGGATCAATCCTGTGATGTGAGAAAGCTTGTGAG 181
DB |||||
QY 600 CTATGATCAGTCACCTATATGACAGCTTGGATTTTGAACCTGTGCATCTGAAGACC 659
DB |||||
QY 182 CTATGATCAGTCACCTATATGACAGCTTGGATTTTGAACCTGTGCATCTGAAGACC 241
DB |||||
QY 660 AGGGAGCTGATTGCATCTACAAAATTTGACGCTCATTAACAAGTGGCATTTGGATCAGCT 719
DB |||||
QY 242 AGGGAGCTGATTGCATCTACAAAATTTGACGCTCATTAACAAGTGGCATTTGGATCAGCT 301
DB |||||
QY 720 GTTTTACAAGCATAAATTTAGCCGTGTTATCATCTAGAGATGATATGGAATTTGCCCC 779
DB |||||
QY 302 GTTTTACAAGCATAAATTTAGCCGTGTTATCATCTAGAGATGATATGGAATTTGCCGC 361
DB |||||
QY 780 TGATTTTTCAGCTTTTTCAGGCTGGAGCTACTCTTCTTGACAGACAGTCGATTAT 839
DB |||||
QY 362 TGATTTTTCAGCTTTTTCAGGCTGGAGCTACTCTTCTTGACAGACAGTCGATTAT 421
DB |||||
QY 840 GGCTATTTCTTCTTGGAAATGACAATGGACAAATGACAGTTTGTCCAAAGATCCTTATGCTCT 899
DB |||||
QY 422 GGCTATTTCTTCTTGGAAATGACAATGGACAAAGGAGTTCGTCCTCAAGATCCTGATGCTCT 481
DB |||||

```

Qy 900 TTACGGCTCAGATTTTTCCTCCGCTCTTGGATGATGCTTTTCAAAATCTACTTGGGACGA 959
Db 482 TTACCGCTCAGACTTTTTCCTCGTCTTGGATGATGCTTTTCAAAATCAACTTGGTCCGA 541
Qy 960 ATTATCTCCAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTAAGACTCAAGAGAATCA 1019
Db 542 ACTATCTTCAAGTGGCCAAAGGCTTACTGGGATGACTGGCTAAGGCTGAAGAAATCA 601
Qy 1020 CAGAGGTCGACAAATTTATTCGCCCGCAGAAAGTTTGCAGAACAT 1060
Db 602 CAGAGGTCGACAAATTTATTCACCCAGAAAGTTTGCAAAACGT 642

RESULT 8
Al486771 643 bp mRNA linear EST 18-MAY-2001
LOCUS EST245093 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
DEFINITION cLED11D20, mRNA sequence.
ACCESSION Al486771
VERSION Al486771.1 GI:4382142
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 643)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
Holt,I.E., Liang,F., Upton,J., Ronning,C.M., Craven,M.B.,
Fujii,C.Y., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C.,
Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato carpel tissue
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
Location/Qualifiers
1..643
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLED11D20"
/tissue_type="carpel"
/dev_stage="5 days pre-anthesis to 5 days post-anthesis"
/lab_host="XL1-Blue MRF"
/clone_lib="tomato ovary, TAMU"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLED - Tomato Carpel EST Library. Oligodt-primed and
directionally cloned cDNA in vector Lambda ZAP II with 5',
and 3' ends located at the EcoRI and XhoI sites,
respectively."

ORIGIN
Query Match 33.3%; Score 578.6; DB 1; Length 643;
Best Local Similarity 93.9%; Pred. No. 9e-154;
Matches 602; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 420 GTTGATCGGAATGTACAGATGCCAGTGGCTGCTGTAGTTGTATGCGTTGCAATCGGC 479
Db 1 GTTAATCGGAATGTGCAGATGCCAGTGGCTGTAGTTGTATGCGTTGCAATCGTTTC 60
Qy 480 TCATTACCTGGAAAGACATTAAATCCATCTTAAATACCAATATCTGTTGCGTCAA 539
Db 61 TGACTACCTGGGAAGACATTAAATCCATCTTAAATACCAATATCTGTTGCGTCAA 120
Qy 540 ATATCTCTTTTCATATCCAGGATGGATCACATCTGTATGTCAGGAAGCTTGTTC 599
Db 121 ATATCTCTTTTCATATCCAGGATGGATCGATCTGTATGATAGAAAGCTTGTTC 180

```

```

Qy 600 CTATGATCAGCTACGCTATATGACGACCTTGGATTTTGAACCTGTGATCTGAAGACC 659
Db 181 CTATGATCAACTGACCTATATGACGACCTTGGATTTTGAACCTGTGATCTGAAGACC 240
Qy 660 AGGGAGCTGATTTGCATCTACAAAATTGCACCTCATTAACAAGTGGCATTGGATCAGCT 719
Db 241 AGGGAACTGGTTGCATCTACCAAGATTGCACGCTATTACAAGTGGCATTGGATCAGCT 300
Qy 720 GTTTTACAAGCATAAATTTTAGCCGCTGTATCATACTAGAAGATGATATGGAATTTGCCCC 779
Db 301 GTTTTACAAGCATAAATTTTAGCCGCTGTATCATACTAGAAGATGATATGGAATTTGCCGC 360
Qy 780 TGATTTTTTTGACTTTTTTGGAGCTGGAGCTACTCTTCTTGAACAGACAGTCGATTAT 839
Db 361 TGATTTTTTTGACTACTTTTGGAGCTGGAGCTACTCTTCTTGAACAGACAGTCGATTAT 420
Qy 840 GGCTATTTCTTCTTGGAGTACAAATGACAAATSCAGTTTGTCTCAAGATCTTATGCTCT 899
Db 421 GGCTATTTCTTCTTGGAGTACAAATGACAAATGACAAATGCTTTCAAAATCAACTTGTCTCT 480
Qy 900 TTACCGCTCAGATTTTTCCTCCGCTCTTGGATGATGCTTTCAAAATCTACTTGGGACGA 959
Db 481 TTACCGCTCAGACTTTTTCCTCGTCTTGGATGATGCTTTCAAAATCAACTTGTCTCT 540
Qy 960 ATTATCTCCAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTAAGACTCAAGAGAATCA 1019
Db 541 ACTATCTCCAAAGTGGCCGAGGCTTACTGGGATGACTGGCTGAGGCTCAAGAAATCA 600
Qy 1020 CAGAGGTCGACAAATTTATTCGCCCGCAGAAAGTTTGCAGAACAT 1060
Db 601 CAGAGGTCGACAAATTTATTCGCCCGCAGAAAGTTTGCAGAACAT 641

RESULT 9
Bg889872 661 bp mRNA linear EST 07-MAR-2003
LOCUS EST515723 cSTD Solanum tuberosum cDNA clone cSTD15H10 5' sequence,
DEFINITION mRNA sequence.
ACCESSION Bg889872
VERSION Bg889872.1 GI:14266958
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 661)
van der Hoeven,R., Bezzerides,J., Ewing,E., Cho,J., Chiemiango,A.,
Bougril,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
Generations of ESTs from dormant potato tubers
Unpublished (2001)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13F-R.
Location/Qualifiers
1..661
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTD15H10"
/tissue_type="dormant tuber"
/dev_stage="one month post-harvest"
/lab_host="SOLR"
/clone_lib="cSTD"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; This library targets genes expressed in dormant
tubers. This library was made from sections of dormant

```

tuber, avoiding the buds and epidermis. Tubers were stored for one month post-harvest at 40C. The tuber was peeled, well away from the surface. Then it was chopped into 1-2 mm cubes and immediately frozen in liquid nitrogen. This library is noted as P4 in Tanksley lab notebooks."

ORIGIN

Query Match	31.9%;	Score	553.6;	DB	2;	Length	661;
Best Local Similarity	94.4%;	Pred. No.	1.3e-146;				
Matches	574;	Conservative	0;	Mismatches	34;	Indels	0;
Gaps	0;						
Qy	126	GATGAGAGGGAACAAGT	TTTGTCTGCTGATTTTC	CGGTACTCTCTCATCTT	TGGTCTGCTGCGC	185	
Db	54	GATGAGAGGGAACAAGT	TTTGTCTGATTTACGGT	ACTTCTCTCTGCTGCTG	CTCTCGC	113	
Qy	186	CTTCATCTACACACAGAT	TGGCGCTTTTGGACAC	AGTCAGCAATATGAC	ATGATGACCGCTTGC	245	
Db	114	CTTCATCTACACACAGAT	TGGCGCTTTTGGACAC	AGTCAGCAATATGAC	ATGATGACCGCTTGC	173	
Qy	246	TGCTGCAATTGAACGAG	AAATCATTTGTACAGC	CCAGCCAGATTCCTT	ATTGACCCAGAT	305	
Db	174	TGCTGCAATTGAACGAG	AAATCATTTGTACAGC	CCAGCCAGATTCCTT	ATTGACCAAGAT	233	
Qy	306	TAGCCTGCAGCAAGAA	TAGTTGCTCTTGAAGA	CAAAATGAAGCGT	CAGGACCCAGGA	365	
Db	234	TAGCAGCAGCAAGAA	TAGTTAGCTCTTGAAGA	CAAAATGAAGCGT	CAGGACCCAGGA	293	
Qy	366	GTGCGCAATTAAGGCT	CTTGTTCAGGATCTT	TGAAGTAAGGGC	ATAAAAAAGTTGAT	425	
Db	294	GTGCGCGCAATTAAGG	CTTGTTCAGGATCTT	TGAAGTAAGGGC	ATAAAAAAGTTAAT	353	
Qy	426	CGGAATGTACAGATGC	CGCTGCTGTAGTTGT	TATGCTTGCAATCGG	CTGATTA	485	
Db	354	CGSAGATGTGAGATGC	CGCTGCTGTAGTTGT	TATGCTTGCAATCGG	CTGCTGACTA	413	
Qy	486	CTTGAAAGAACTATT	AAATCCATCTTAAATA	CAAAATATCTGCTG	CGTCAAAATATCC	545	
Db	414	CTTGAGAGGACTATT	AAATCCATCTTAAATA	CAAAATATCTGCTG	CGTCAAAATATCC	473	
Qy	546	TCCTTTTCATATCCAG	GATGGATCACATCTCT	GTATGTTCAGGAAG	CTGCTTTTGAGCTATGA	605	
Db	474	TCCTTTTCATATCCAG	GATGGATCAAATCTCT	GTATGTAAAGAAAG	CTGCTTTTGAGCTATGA	533	
Qy	606	TCAGCTACGCTATATG	ACGACTTGGATTTTGA	ACCTGTGCATCTGA	AAAGACCCAGGGA	665	
Db	534	TCAGCTACGCTATATG	ACGACTTGGATTTTGA	ACCTGTGCATCTGA	AAAGACCCAGGGA	593	
Qy	666	GCTGATTGCATCTACA	AAAAATTGCACGCTCA	TTCAAGTGGGCATTT	GATCAGCTGTTTTA	725	
Db	594	ACTGTTGCATCTACA	AGATTGACGCTCATT	ACAGTGGGCATTTG	ATCAGCTGTTTTCA	653	
Qy	726	CAAGCATA	733				
Db	654	CAAGCATA	661				

RESULT 10

AW979500
LOCUS
DEFINITION
EST310521 tomato root deficiency, Cornell University Lycopersicon
esculentum cDNA clone CLEW209 5', mRNA sequence.

AW979500

AW979500.1 GI:8171019

EST.

Lycopersicon esculentum (Solanum lycopersicum)

ORGANISM

Lycopersicon esculentum
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
Spermatophyta: Magnoliophyta: eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE

1 (bases 1 to 673)
van der Hoeven,R.S., Garvin,D.F., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T.S., Ronning,C.M., Craven,M.B., Bowman,C.L.,
Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Giovannoni,J.J.

and Tanksley,S.D.
Generation of ESTs from tomato nutrient-deficient roots
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

COMMENT

FEATURES

Location/Qualifiers
1..673

source

/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="GLEW209"
/issue_type="roots"
/dev_stage="5-6 weeks old"
/clone_lib="tomato root deficiency, Cornell University"
/notes="Vector: pBluescript SK-; Site_1: 5' EcoRI; Site_2:
3' XhoI; supplier: Tanksley; Tissue supplied by Dave
Garvin (USDA-ARS, Ithaca, NY 14850). Roots were harvested
from plants grown under the following
deficiencies/stresses: 10 mM Al, Zn, P, K, Fe, N), and
mRNA was isolated from individual treatments. Proportional
aliquots of mRNA of each treatment were mixed and used for
library construction."

ORIGIN

Query Match	31.0%;	Score	537.8;	DB	1;	Length	673;
Best Local Similarity	93.1%;	Pred. No.	4.4e-142;				
Matches	563;	Conservative	0;	Mismatches	42;	Indels	0;
Gaps	0;						
Qy	126	GATGAGAGGGAACAAGT	TTTGTCTGCTGATTT	TCGGTACCTCTCTCATCT	TGTGCTGCTGTCG	185	
Db	69	GATGAGAGGGAACAAGT	TTTGTCTGATTTACGGT	ACCTTCTCTGCTGCTG	CTCTCGC	128	
Qy	186	CTTCATCTACACACAGAT	TGGCGCTTTTGGACAC	AGTCAGCAATATGCA	GATGCCTTGC	245	
Db	129	CTTCATCTACACACAGAT	TGGCGCTTTTTCGTGAC	ACAGTCAGATATGCA	GACGCGCTTGC	188	
Qy	246	TGCTGCAATTGAACGAG	AAATCATTTGTACAGC	CCAGCCAGATTCCTT	ATTGACCCAGAT	305	
Db	189	TGCTGCAATTGAACGAG	AAATCATTTGTACAGC	CCAGCTAGCTTCTT	ATTGACCAAGAT	248	
Qy	306	TAGCCTGCAGCAAGAA	GAATAGTCTCTTGAAGA	CAAAATGAAGCGT	CAGGACCCAGGA	365	
Db	249	TAGCCAGCAGCAAGAA	GAGTAGTGGCTCTTGAAGA	CAAAATGAAGCGC	CAGGACCCAGGA	308	
Qy	366	GTGCGCACAATTGAAGG	CTCTTGTTCAGGATCTT	TGAAGTAAGGGCAT	TAATAAAGTTGAT	425	
Db	309	GTGCGCACAAATTGAAG	GCTCTTGTTCAGGATCTT	TGAAGTAAGGGCAT	TAATAAAGTTAAT	368	
Qy	426	CGGAATGTACAGATGC	CGCTGCTGTAGTTGT	TATGGCTTGAATCGG	CTGATTA	485	
Db	369	CGGAATGTGAGATGCC	AGTCCAGCTGTAGTTG	TATGGCTTGAATCGG	CTGCTGACTA	428	
Qy	486	CTTGAAAAAGACTATT	AAAAATCCATCTTAAAA	TACCAAAATATCTGT	TGCGTCAAAATATCC	545	
Db	429	CTTGAGAGAGACTATA	AAAAATCCATCTTAAAA	TACCAAAATATCTGT	TGCGTCAAAATATCC	488	
Qy	546	TCCTTTTCATATCCAGG	ATGGATCATCTCTGAT	GTGCAGGAAGCTTGT	TGAGCTATGA	605	
Db	489	TCCTTTTCATATCCAGG	ATGGATCGAATCCTGT	ATGTAGAAAGCTTGT	TGAGCTATGA	548	
Qy	606	TCAGCTGACGTATATC	CAGCCTTGGATTTTGA	ACCTGTGCATCTGA	AGAACACAGGGA	665	
Db	549	TCACTGAGGTATATGC	AGCATTGGATTTTGA	ACCTGTGCATCTGA	AGAACACAGGGA	608	
Qy	666	GCTGATTGCATCTACA	AAAAATTGCAGCTCAT	TACAAAGTGGGCAT	TGGATTCAGCTGTTTA	725	
Db	609	ACTGGTTGCATCTACA	AGATTGCAGCTCAT	TACAAAGTGGGCAT	TGGATTCAGCTGTTCA	668	

```

Oy 726 CAAGC 730
    |||||
Db 669 CAAGC 673

RESULT 11
AW735827      659 bp      mRNA      linear      EST 18-MAY-2001
LOCUS      EST33595 tomato flower buds 0-3 mm, Cornell University
DEFINITION  Lycopersicon esculentum cDNA clone c10A5111 5', mRNA sequence.
ACCESSION  AW735827
VERSION    AW735827.1 GI:7642666
KEYWORDS   EST.
SOURCE     Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM   Lycopersicon esculentum
REFERENCE   1 (bases 1 to 659)
AUTHORS    van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E.,
            Liang,F., Hansen,T., Craven,M.B., Bowman,C.B., Romning,C.M.,
            Niernan,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and
            Tanksley,S.D.
            Generation of ESTs from tomato flower tissue, 0-3 mm buds
            Unpublished (1999)
            Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
            Location/Qualifiers
            1..659
            /organism="Lycopersicon esculentum"
            /mol_type="mRNA"
            /cultivar="TA496"
            /db_xref="taxon:4081"
            /clone="CT0A5111"
            /tissue_type="Flower"
            /dev_stage="0-3mm buds"
            /clone_lib="tomato flower buds 0-3 mm, Cornell University"
            /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
            XhoI; supplier: Tanksley; Flower buds and flowers were
            taken from greenhouse plants (4-8 wks old, TA496). They
            were immediately frozen in liquid nitrogen and then
            size-separated while remaining frozen."

FEATURES
    source
    1..659
    /organism="Lycopersicon esculentum"
    /mol_type="mRNA"
    /cultivar="TA496"
    /db_xref="taxon:4081"
    /clone="CT0A5111"
    /tissue_type="Flower"
    /dev_stage="0-3mm buds"
    /clone_lib="tomato flower buds 0-3 mm, Cornell University"
    /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
    XhoI; supplier: Tanksley; Flower buds and flowers were
    taken from greenhouse plants (4-8 wks old, TA496). They
    were immediately frozen in liquid nitrogen and then
    size-separated while remaining frozen."

ORIGIN
Query Match      30.3%; Score 527; DB 1; Length 659;
Best Local Similarity 93.2%; Pred. No. 5.3e-139;
Matches 551; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Oy 126 GATGAGAGGGAACAAGTTTGTGTTATTCGGTACCTCCATCTTGGCTGCTGTCG 185
    |||||
Db 69  GATGAGAGGAACAAGTTTGTGTTATTCGGTACCTCCATCTTGGCTGCTGTCG 128
    |||||

Oy 186 CTTTCATCTACACAGATCGGCTTTTTCGCACAGTCAGATATGCAGATCGCCTTGC 245
    |||||
Db 129 CTTTCATCTACATACAGATCGGCTTTTTCGTGACACAGTCAGATATGCAGATCGCCTTGC 188
    |||||

Oy 246 TGTGCAATTAAGCAGAGAAATCAATTTGACAGCCAGACAGATTCCTATTGACCCAGAT 305
    |||||
Db 189 TGTGCGATTGAAGCAGAGAAATCAATTTGACAGTCAGATAGGTTCCTATTGACCAAGAT 248
    |||||

Oy 306 TAGCCTGCAGCAGGAGGAATAGTCTCTTTGAAGAACAAATGAAGCGTCAGGACCCAGGA 365
    |||||
Db 249 TAGCCAGCAGCAGGAGGAGTAGTGGCTCTTGAAGAACAAATGAAGCGGCGGACCCAGGA 308
    |||||

Oy 366 GTCCCGACAAATTAAGGGCTCTTCTTCAGGATCTTTGAAGTAAGGGCATAAATAAGTTGAT 425
    |||||
Db 309 GTGCCAGCAATTAAGGGCTCTTCTTCAGGATCTTTGAAGTAAGGGCATAAATAAGTTAAT 368
    |||||
Oy 426 CGGAATGTACAGATGCCAGTGGCTGCTGTAGTTGTTATGGCTTGCAATCGGCTGATTA 485
    |||||

```

```

Db 369 CGGAATGTGCAGATGCCAGTGGCTGCTGTAGTTGTTATGGCTTGCAATCGGCTGATTA 428
    |||||
Oy 486 CCTGGAAGAAAGACTATTAAATCCATCTTAAATACCAATATCTGTTGCGTCAAAATATCC 545
    |||||
Db 429 CCTGGAGAAGACTATAAAATCCATCTTAAATACCAATATCTGTTGCAATCAAAATATCC 488
    |||||
Oy 546 TCTTTTCATATCCAGGATGGATCACATCTCTGATGTGAGGAAGCTTGTCTTTGAGCTATGA 605
    |||||
Db 489 TCTTTTCATATCCAGGATGGATCGAATCTGATGTGAGGAAGCTTGTCTTTGAGCTATGA 548
    |||||
Oy 606 TCAGCTGACGTATATGAGCAGCACTTGGATTTTGAACCTGTGCATCTGAAACACGAGGGA 665
    |||||
Db 549 TCAACTGACGTATATGAGCAGCACTTGGATTTTGAACCTGTGCATCTGAAACACGAGGGA 608
    |||||
Oy 666 GCTGATTGCATCTACTACAAAATTTGCACGTCTATTCAAGTGGCATTGGATCA 716
    |||||
Db 609 ACTGTTGCATCTACTACAAAATTTGCACGTCTATTCAAGTGGCATTGGATCA 659
    |||||

RESULT 12
BM779745      843 bp      mRNA      linear      EST 04-MAR-2002
LOCUS      EST590321 KV2 Medicago truncatula cDNA clone PKV2-52N22, mRNA
DEFINITION  sequence.
ACCESSION  BM779745
VERSION    BM779745.1 GI:19109675
KEYWORDS   EST.
SOURCE     Medicago truncatula (barrel medic)
ORGANISM   Medicago truncatula
REFERENCE   1 (bases 1 to 843)
AUTHORS    VandenBosch,K., Endre,G., Hur,J., Beremand,P., Town,C.D., Van
            Aken,S., Utterback,T., Cheung,F., Tsai,J. and Fraser,C.M.
            ESTs from roots of Medicago truncatula 48 hr after inoculation with
            Sinorhizobium meliloti
            Unpublished (2002)
            Contact: VandenBosch K
            Department of Plant Biology
            University of Minnesota
            220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
            Tel: 612 624 2755
            Fax: 612 625 1738
            Email: kvandenb@cbs.umn.edu
            TIGR sequence name: MTACP83TK More information is available at:
            www.medicago.org
            Seq primer: SKmod (CTA GAA CTA gtg gat CC) .
            Location/Qualifiers
            1..843
            /organism="Medicago truncatula"
            /mol_type="mRNA"
            /cultivar="genotype A17"
            /db_xref="taxon:3880"
            /clone="pKV2-52N22"
            /tissue_type="Seedling roots"
            /dev_stage="2 days post-inoculation with Sinorhizobium
            meliloti"
            /lab_host="E. coli strain SOLR"
            /clone_lib="KV2"
            /note="Vector: pBluescript SK -; Site 1: EcoRI; Site 2:
            XhoI; cDNA was prepared from polyA+ enriched RNA. The
            cDNA was directionally ligated into the Unipax XR vector
            from Stratagene and packaged using Gigapack III Gold
            packaging extracts. Plasmids containing cDNA inserts
            were excised from the recombinant lambda-Zap phage using
            Ex-assist helper phage and propagated in SOLR cells."

ORIGIN
Query Match      29.8%; Score 517.2; DB 3; Length 843;
Best Local Similarity 75.9%; Pred. No. 3.7e-136;

```


Db 484 NTCTGCTTGGATGGATGCTTTCAAATCAACTTGGTCCGAACATATCTCCAAAGTGG 543

Qy 976 CCAAGGCTTACTGGGACGACTGGCTAAGAC 1006

Db 544 CCGAAGGCTTACTGGGATGACTGGCTGAGC 574

RESULT 14

AJ806862 758 bp mRNA linear EST 08-DEC-2004

LOCUS AJ806862 Antirrhinum majus whole plant Antirrhinum majus cDNA clone

DEFINITION 018 6 05 c08, mRNA sequence.

ACCESSION AJ806862

VERSION AJ806862.1 GI:51122190

KEYWORDS EST.

ORGANISM Antirrhinum majus (snapdragon)

Antirrhinum majus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiales; Plantaginales; Antirrhineae; Antirrhinum.

REFERENCE 1 (bases 1 to 758)

AUTHORS Bey, M., Stueber, K., Fellenberg, K., Schwarz-Sommer, Z., Sommer, H., Siedler, H. and Zachgo, S.

TITLE Characterization of Antirrhinum Petal Development and Identification of Target Genes of the Class B MADS Box Gene

JOURNAL DEFICIENS

PUBMED Plant Cell 16 (12), 3197-3215 (2004)

COMMENT 15539471

Contact: Schwarz-Sommer Z

Molekulare Pflanzen-genetik

MPI fuer Zuechtungs-forschung

Carl-von-Linne Weg 10, D-50829, Germany.

Location/Qualifiers

1..758

/organism="Antirrhinum majus"

/mol_type="mRNA"

/db_xref="taxon:4151"

/clone="018 6 05 c08"

/tissue_type="whole plant"

/clone_lib="Antirrhinum majus whole plant"

ORIGIN

Query Match 29.4%; Score 510.6; DB 1; Length 758;

Best Local Similarity 82.0%; Pred. No. 2.8e-134;

Matches 588; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

Qy 552 CATATCCAGGATGGATCACATCCCTGATGTCAGGAAGCTTGTCTTGAGCTATGATCAGCT 611

Db 42 CGTATCACAGATGGATCAAAATCTGATGTTAAAGCATGGCGTTGAGTTATGATCAGCT 101

Qy 612 GAGCTATATGACGACTTGGATTTTGAACCTGTGCATCTGAAAGACCGAGGAGCTGAT 671

Db 102 AACCTACATGCAGCATTTGGACTATGACAGATCCGTACAGAAAGCGCTGGCGAGTTGAT 161

Qy 672 TGCATACTACAAAATTTGACGCTATTACAGTGGGCAATGGATCAGCTTTTACAGCA 731

Db 162 CGCATACTACAAGATGACGCGCATTTACAAATGGGCACTTGATCAATTTTACAAACA 221

Qy 732 TAATTTAGCCGTGTATCATCTAGACATGATAGAAATGGCCCTGATTTTTCGA 791

Db 222 TAAATTTAGCCGATGATATCATCTCGAAGATGATAGAAATGGCCCTGATTTCTTGA 281

Qy 792 CTTTTTTAGGCTGGAGCTACTCTCTTCACAGACAAAGTCGATTTATGGCTATTTCTTC 851

Db 282 GTATTTTGGCGGAGCGGCACTTTTCACCGTGACAATCCATCATGCAATTTTCATC 341

Qy 852 TTGAATGATCAATGGACAAATGACGTTTGTCCAAAGATCTTATGCTCTTTACCGCTCAGA 911

Db 342 ATGGAATGATATGGACAAACGCGAGTTTGTACATGATCCTTATGCTGTACCGCTCAGA 401

Qy 912 TTTTTCCTCCGCTCTGGATGGATGCTTTCAAATCTACTTGGGACGAATATCTCCAAA 971

Db 402 TTTTTCCTCCGCTCTGGATGGATGCTTTCAAATCAACTGCGGATGAAATGTCTCTCTAA 461

Qy 972 GTGSCCAAAGGCTTACTTGGGACGACTGGCTAAGACTCAAAGAGAAATCACAGAGTCCGACA 1031

Db 462 ATGCCCCAAAGCTTACTTGGGATGACTGTTGCGACTCAAGAGAAATCACAAAGGACGGCA 521

Qy 1032 ATTTATTCGCCCAAGCTTTTCAGAACATATAATTTTGGTGAGCATGGTCTTAGTTTGGG 1091

Db 522 ATTCATTCGCCCAAGCTATGACAGAACATATAATTTTGGGAGCATGGATCCAGCATGGG 581

Qy 1092 GCAGTTTTTCAAGCAGTATCTTGAGCCAAATTAACCTAAATGATCTCCAGGTTGATTGAA 1151

Db 582 GCAATTTTCAACAGATATCTAGAACCTATCAAGATGAATGATATCCAGGTTGATTGAA 641

Qy 1152 GTCAATGGACCTTAGTTTACCTTTTGGAGGACAAATTAAGTCTGAAACACTTTGGTGACTTGGT 1211

Db 642 ATCAATGAGCTCAGCTATCTGGAGGAGGACAAAGTATGTGAAGCACITTTGCTGACTTGT 701

Qy 1212 TAAAAAGGCTAAGCCCATCCATGAGGATGCTGTGCTTGAAGCAATTTAAACATAGA 1268

Db 702 TAAAAAGGCGAGCGCTGTGCATGGGCTGCTGCTCTGAAAGGCAAAATATATAGA 758

RESULT 15

BM780056

LOCUS BM780056 852 bp mRNA linear EST 04-MAR-2002

DEFINITION EST590632 KV2 Medicago truncatula cDNA clone pkV2-53H18, mRNA

sequence.

ACCESSION BM780056

VERSION BM780056.1 GI:19110092

KEYWORDS EST.

SOURCE Medicago truncatula (barrel medic)

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE 1 (bases 1 to 852)

AUTHORS VandenBosch, K., Endre, G., Hur, J., Beremand, P., Town, C.D., Van Aken, S., Utterback, T., Cheung, F., Tsai, J. and Frazer, C.M.

TITLE ESTs from roots of Medicago truncatula 48 hr after inoculation with Sinorhizobium meliloti

JOURNAL Unpublished (2002)

COMMENT Contact: VandenBosch K

Department of Plant Biology

University of Minnesota

220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA

Tel: 612 624 2755

Fax: 612 625 1738

Email: kvandenb@cbs.umn.edu

TIGR sequence name: MTACT45TK More information is available at: www.medicago.org

Seq primer: SKmod (CTA GAA CTA gtg gAT CC).

Location/Qualifiers

1..852

/organism="Medicago truncatula"

/mol_type="mRNA"

/cultivar="genotype A17"

/db_xref="taxon:3880"

/clone="pkV2-53H18"

/tissue_type="Seedling roots"

/dev_stage="2 days post-inoculation with Sinorhizobium meliloti"

/lab_host="E. coli strain SOLR"

/clone_lib="KV2"

/note="Vector: pBluescript SK -; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-zap phage using Ex-assist helper phage and propagated in SOLR cells."

ORIGIN

Query Match		29.3%;	Score 508.2;	DB 3;	Length 852;
Best Local Similarity		75.3%;	Pred. No. 1.4e-133;		
Matches 633;		Conservative 0;	Mismatches 208;	Indels 0;	Gaps 0;
Qy	142	TTTTGCTGTGATTTCCGGTACCTCTCATCTTGGCTGCTGTGCGCTTTCATCTACACACAG	201		
Db	2	TTTTGGTGTGATTTTCGGTTCCTCTCTTTGTCAGCCITGGTGTTCATCTACATCCAG	61		
Qy	202	ATGGCGCTTTTGGACACAGTCAGAATATGCAGATCGCCTTGCTGTGCTGCAATTTGAAGCA	261		
Db	62	ATGAGACTTTTCGCATCACAAATACAATATGCTGATCGCCTCGCTGCTGTATTGAAGCT	121		
Qy	262	GAAATCATTTGTACAAGCCAGACAGATTTGCTTATATGACCAAGATTAGCCTGCACAGGA	321		
Db	122	GAAATCACTGTACAGCTCAAAATCCGATCACTATTGATCAGATTAGCTTGCAACAAGGA	181		
Qy	322	AGATAGTTGCTCTTTGAAGAACAAATGAAGCGTCAGGACCGAGAGTCCGCACAATTAAGG	381		
Db	182	CGCATTCGACTTACAAACGAAGAAGAACCGTCGAGAGCAAGAATGTAGCCNAATCAAG	241		
Qy	382	GCTCTTGTTCAGGATCTTTGAAGTAAGGGCATAAAAAAGTTGATCGGAAATGTACAGATG	441		
Db	242	TCTCTTGTTCAGATCTTTGAAGAAAGAGATGTAAGAGGCTGATTTGATAAAGTCCAGGTT	301		
Qy	442	CCAGTGGCTGCTGTAGTTGTTATGGCTTGCAATCGGGCTGATTAACCTGGAAAAAGACTATT	501		
Db	302	CCAGTGGCAGCTGTGTGCATCATGGCATGTAATCGTGTGATTACCTGGAGAGGACTATC	361		
Qy	502	AAATCCATCTTAAATACCAAAATATCTGTGCGTCAAAATATCCTCTTTTCATATCCCG	561		
Db	362	AATTCTGTATTGAATATCAAGGCCCATTTCTTCAAGATTTCTTTATTGTATCACAG	421		
Qy	562	GATGGATCACATCTGTATGTCAGGAAGCTTGCTTTGAGCTATGATCAGCTGACGCTATATG	621		
Db	422	GATGGATCAAAATTCAGACGTCAAAAGAAAGGCTCTGAGCTACGATGAGCTATCTCATATG	481		
Qy	622	CAGCATTGGATTTTGAACCTGTGCATCTGAAGAACAAGGGAGCTGATTCATACATAC	681		
Db	482	CAGCATTAGATTTTGAACCGGTTCAAACTGAAAGGCCAGGAGAGTTAAATTTGCTTACTAT	541		
Qy	682	AAATTTGCACGTCATTACAAGTGGCATTTGATCAGCTGTGTTTACAAGCATAAATTTAGC	741		
Db	542	AAGATTGCACGTCATTACAAGTGGCTCTTGGTCAACTGTTTCGACAAAGCATAAATTTCAAC	601		
Qy	742	CGTGTATCATCTAGAAGATGATATGGAAATTCGCCCTGATTTTGTGACTTTTGTGAG	801		
Db	602	AGAGTCATCATTTCTGAAGATGACATGGAATAGCACCTGATTTCTTTGATTATTTTGAA	661		
Qy	802	GCTGGAGCTACTCTTTCTGACAGACAGAAAGTCGATATGCGTATTTCTTTTGGAAATGAC	861		
Db	662	GCTATGGCAACTCTCTTCAGCAGGATAAGTCCATTATGCGCTGTTTCGTATGGAATGAC	721		
Qy	862	AATGGACAAATGCAAGTTTGTCCAAAGATCCTTATGCTCTTTACCGCTCAGATTTTTTCCC	921		
Db	722	AATGGACAGAAAGCAGTTTGTACATGATCCATATGAACCTTATCGCTCGGACTTCTTTCCCT	781		
Qy	922	GGTCTTGGATGGATGCTTTTCAAAATCTACTTTGGGACGAATTATCTCCAAAGTGCCAAAG	981		
Db	782	GGACTAGGATGGATGCTGGCTAGATCTACATGGGATGAACTATCACCAAAATGSCCCGAA	841		
Qy	982	G 982			
Db	842	G 842			

Search completed: December 14, 2005, 13:02:49
Job time : 6402 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 06:18:00 ; Search time 8626 Seconds
(without alignments)
11446.450 Million cell updates/sec

Title: US-10-615-144-3

Perfect score: 1737
Sequence: 1 gaattcgccgcgcattga.....ataatcgccgcgcgaattc 1737

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb.ba.*
2: gb.in.*
3: gb.env.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pr.*
9: gb.ro.*
10: gb.sts.*
11: gb.sy.*
12: gb.un.*
13: gb.vi.*
14: gb.hcg.*
15: gb.pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1737	100.0	1737	6	A95055 Sequence 3
2	1737	100.0	1737	6	BD226905 Plant Gnt
3	1737	100.0	1737	6	AR432619 Sequence
4	1708	98.3	1708	15	NTA249883 Nicotiana
5	1561.8	89.9	2031	15	Y16832 Nicotiana t
6	1492.6	85.9	1836	15	AJ249882 Nicotiana
7	1297.8	74.7	1341	15	NBE295993 Nicotiana
8	1255	72.3	1641	15	STU249878 Solanum t
9	1255	72.3	1669	6	A95053 Sequence 1
10	1255	72.3	1669	6	BD226904 Plant Gnt
11	1255	72.3	1669	6	AR432618 Sequence
12	1239.2	71.3	1719	15	BT013875 Lycopersi
13	1216.6	70.0	1691	15	STU249879 Solanum t
14	1078	62.1	1499	15	AJ249880 Solanum t
15	772	44.4	1625	15	BT000334 Arabidops
16	772	44.4	1766	15	AY099838 Arabidops
17	770.6	44.4	1820	15	AJ243198 Arabidops
18	767.2	44.2	1830	15	AJ249881 Arabidops

19	767.2	44.2	1854	6	A95057
20	767.2	44.2	1854	6	BD226906
21	767.2	44.2	1854	6	AR432620
22	688.2	38.5	1329	15	OS4457976
23	688.2	38.5	1897	15	AK101526
24	658.6	37.9	2180	6	AX840670
25	658.6	37.9	9240	6	AX840669
26	658	37.9	1319	6	AX840672
27	656.6	37.8	1119	6	AX840688
28	425.8	24.5	896	15	AK058826
29	419.6	24.2	1415	15	PPA429143
30	417.2	24.0	422	15	NTA295994
31	183.4	10.6	2695	9	MUSTRANSA
32	181.8	10.5	1894	9	MUSGLCNACT
33	181.8	10.5	2651	9	BC006629
34	181.8	10.5	203286	9	AL606829
35	181.6	10.5	2545	9	RATNAGT
36	181.6	10.5	2557	6	E07296
37	181.6	10.5	2595	9	BC074010
38	180.2	10.4	2988	9	BC031752
39	173.6	10.0	1344	9	AF087456
40	172.8	9.9	2057	5	BC088510
41	172	9.9	1344	9	AF087457
42	165.8	9.5	1519	2	AK115237
43	165.8	9.5	2894	5	XLGLCNAC
44	164.6	9.5	1344	9	CGU65792
45	164.2	9.5	2243	5	BC041180

ALIGNMENTS

RESULT 1

A95055 1737 bp DNA linear PAT 26-JAN-2000
Sequence 3 from Patent WO929879.
A95055
ACCESSION A95055.1 GI:6779208
KEYWORDS
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE

1 (bases 1 to 1737)
Von,S.A.

AUTHORS

TITLE

JOURNAL

FEATURES

source

Location/Qualifiers

1..1737

/organism="Nicotiana tabacum"

/mol_type="unassigned DNA"

/cultivar="Samsun NN"

/db_xref="taxon:4097"

/cell_type="BLATTZELLEN"

/tissue_type="MESOPHYLL"

/clone_lib="LAMBDA ZAP II (ECORI)"

/dev_stage="M: 'SINK' ORGAN"

15..126

127..1467

/gene="CGL"

/gene="CGL"

/EC_number="2.4.1.101"

/function="INITIIERT KOMPLEXE N-GLYKANE AUF SEKRET."

GLYKOPROTEINEN"

/standard_name="GNTI"

/note="ERSTE GNTI-SEQUENZ AUS TABAK (UNPUBLIZIERT)"

/codon_start=1

/evidence=experimental

5'UTR

gene

CDS

		/product="BETA-1,2-N-ACETYLGLUCOSAMINYLTRANSFERASE I"	
		/protein_id="CAB70464.1"	
		/db_xref="GI:6779209"	
		/translation="MRGNKFCDDFRYLLILAAVAFIYQMLRFATSEYADRLAAALIE	
		AENHCTGOTRLIDIOISLQOGRIVALEEQMKRODOECQRLRALVQDLSEKGIKKLIGN	
		VOMPVAAVVMACNRADYLEKTIKILKYOISVASKYPLFISODGSHPDVVKLALSVD	
		OLTYMHLDPVHTERPEGLIAYKTKARYKHALDQLFYKHNFSRVILIEDDMETAP	
		DFFDFBAGATLLDRDKSISMAISSWNGNQMFQVDYALYRSDFPFLGLWMLSKSTW	
		DEUSPKWPKAYDMDLRLKENHRQPIRPEVCRTYNFSGHSLGQFFKYLEPIKL	
		NDVQDWKSDLSLYLHGFDLVKKAKPIHGADAVLKAFNIDGDVRIQYRDQL	
		DFENIAKQFGIIEEWKQVPRAAYKGIWFRYQTSRRVFLVGHDSLQQLGIEDT"	
		<154..>213	
CDS		/gene="CGL"	
		/function="MEMBRANANKER (AMINO ACID 10-29)"	
		/standard_name="MEMBRANANKER EINES TYP II GOLGI-PROTEINS"	
		/note="DURCH VERGLEICH MIT TIERISCHEN GNTI-SEQUENZEN	
		IDENTIFIZIERT"	
		/codon_start=1	
		/product="HYDROPHOBER AMINO ACIDBEREICH IN GNTI"	
		/protein_id="CAB70465.1"	
		/db_xref="GI:6779210"	
		/translation="FRYLLILAAVAFIYQMLRF"	
		1468..1723	
3'UTR			
ORIGIN			
Query Match		100.0%; Score 1737; DB 6; Length 1737;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 1737; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	GAATTCGGCGCGCCATTGACTTCATCTCACTCAACAGCAGCAAGTAATCCAGCGATGA	60
Db	1	GAATTCGGCGCGCCATTGACTTCATCTCACTCAACAGCAGCAAGTAATCCAGCGATGA	60
Qy	61	AACACTCACTACTGAACTGAGAGACTATTCGCTTCTCTCTAAAGCCTTCAATCGGAATT	120
Db	61	AACACTCACTACTGAACTGAGAGACTATTCGCTTCTCTCTAAAGCCTTCAATCGGAATT	120
Qy	121	CGCAGATGAGAGGGAACAAGTTTGTGTGTGATTTCCGGTACCTTCCTCATCTTGGCTGCT	180
Db	121	CGCAGATGAGAGGGAACAAGTTTGTGTGTGATTTCCGGTACCTTCCTCATCTTGGCTGCT	180
Qy	181	GTCCGCTTCATCTACACAGATCGGCTTTTTCGGACAGCTCAGATGCAATATGCAATCGC	240
Db	181	GTCCGCTTCATCTACACAGATCGGCTTTTTCGGACAGCTCAGATGCAATATGCAATCGC	240
Qy	241	CTTGCTGCTGCAATTGAAGCAGAAAATCATTTGTACAAAGCCAGACCAAGTTGCTTATTGAC	300
Db	241	CTTGCTGCTGCAATTGAAGCAGAAAATCATTTGTACAAAGCCAGACCAAGTTGCTTATTGAC	300
Qy	301	CAGATTAGCTTCGACGCAAGGAATATGTTGCTCTTGAAGAACCAATGAAGCGTCAGGAC	360
Db	301	CAGATTAGCTTCGACGCAAGGAATATGTTGCTCTTGAAGAACCAATGAAGCGTCAGGAC	360
Qy	361	CAGAGTGCCGACAAATTAAGGGCTCTTGTTCAGAGATCTTGAAGTAAGGGCATAAAAG	420
Db	361	CAGAGTGCCGACAAATTAAGGGCTCTTGTTCAGAGATCTTGAAGTAAGGGCATAAAAG	420
Qy	421	TTGATCGGAAATGACAGATCCAGTGGCTGCTGTAGTTGTTATGGCTTGCATTCGGGCT	480
Db	421	TTGATCGGAAATGACAGATCCAGTGGCTGCTGTAGTTGTTATGGCTTGCATTCGGGCT	480
Qy	481	GATTACCTGGAAAAGACTATTAATTCATCTTAAATAACCAAAATATCTGTTCGCTCAAAA	540
Db	481	GATTACCTGGAAAAGACTATTAATTCATCTTAAATAACCAAAATATCTGTTCGCTCAAAA	540
Qy	541	TATCCTCTTTTCATATCCAGGATGGATCAATCTGATGTCAGGAAGCTTGTCTTCAGC	600
Db	541	TATCCTCTTTTCATATCCAGGATGGATCAATCTGATGTCAGGAAGCTTGTCTTCAGC	600
Qy	601	TATGATCAGCTGATATATGACACCTTGATTTTGAACCTGTGCACTACTGAAAGACCA	660
Db	601	TATGATCAGCTGATATATGACACCTTGATTTTGAACCTGTGCACTACTGAAAGACCA	660

Qy	661	GGGAGCTGATTCGCATACATACTACAAAATTGCACGTCATTACAAGTGGGCATTGGATCAGCTG	720
Db	661	GGGAGCTGATTCGCATACATACTACAAAATTGCACGTCATTACAAGTGGGCATTGGATCAGCTG	720
Qy	721	TTTTCACAGCATAAATTTTAGCCGCTTATCATACATAGAGATGATATGGAATTTGCCCT	780
Db	721	TTTTCACAGCATAAATTTTAGCCGCTTATCATACATAGAGATGATATGGAATTTGCCCT	780
Qy	781	GATTTTGTGACTTTTTTCAGGCTGAGCTACTCTTCTTGACAGACAAGTCGATATG	840
Db	781	GATTTTGTGACTTTTTTCAGGCTGAGCTACTCTTCTTGACAGACAAGTCGATATG	840
Qy	841	GCTATTCTCTCTGGAAATGACAAATGCGAGTTTGTCCAAAGTCCTTATGCTCTT	900
Db	841	GCTATTCTCTCTGGAAATGACAAATGCGAGTTTGTCCAAAGTCCTTATGCTCTT	900
Qy	901	TACCGCTCAGATTTTTTCCCGGCTTGGATGATGCTTTCAAAATCTACTTCGGACGAA	960
Db	901	TACCGCTCAGATTTTTTCCCGGCTTGGATGATGCTTTCAAAATCTACTTCGGACGAA	960
Qy	961	TTATCTCCAAAGTGGCCAAAGGCTTACTGGACGACTGGCTAAAGACTCAAAAGAAATCAC	1020
Db	961	TTATCTCCAAAGTGGCCAAAGGCTTACTGGACGACTGGCTAAAGACTCAAAAGAAATCAC	1020
Qy	1021	AGAGGTGCAACNATTTATTCGCCCAGAGTTTGCAGAACATATATATTTTGGTGAGCATGGT	1080
Db	1021	AGAGGTGCAACNATTTATTCGCCCAGAGTTTGCAGAACATATATATTTTGGTGAGCATGGT	1080
Qy	1081	TCTAGTTTGGGGCAGTTTTTCAAGCAGTATCTTGAGCCAAATTAACCTAAATGATGTCAG	1140
Db	1081	TCTAGTTTGGGGCAGTTTTTCAAGCAGTATCTTGAGCCAAATTAACCTAAATGATGTCAG	1140
Qy	1141	GTTGATTGGAAGTCAAATGGAACCTTACTTAACCTTTTCGAGAGCAATTAACCTGAAACACTTT	1200
Db	1141	GTTGATTGGAAGTCAAATGGAACCTTACTTAACCTTTTCGAGAGCAATTAACCTGAAACACTTT	1200
Qy	1201	GGTGACTTGGTTAAAAGGCTAAGCCCATCCATCCATCCATCCATCCATCCATCCATCCATCCAT	1260
Db	1201	GGTGACTTGGTTAAAAGGCTAAGCCCATCCATCCATCCATCCATCCATCCATCCATCCATCCAT	1260
Qy	1261	AACATAGATGGTGTGATGCGGTATTCAGTACAGAGATCAACTAGACTTTGAAAATATCGCA	1320
Db	1261	AACATAGATGGTGTGATGCGGTATTCAGTACAGAGATCAACTAGACTTTGAAAATATCGCA	1320
Qy	1321	CGGCAATTTGGCAATTTTGAAGAAATGGAAGGATGGTGTACCACTGCAGCATATAAAGGA	1380
Db	1321	CGGCAATTTGGCAATTTTGAAGAAATGGAAGGATGGTGTACCACTGCAGCATATAAAGGA	1380
Qy	1381	ATAGTAGTTTTCGGGTACCAACGTCAGACGTTATTCCTTGTTCGGCCATGATTCGCTTT	1440
Db	1381	ATAGTAGTTTTCGGGTACCAACGTCAGACGTTATTCCTTGTTCGGCCATGATTCGCTTT	1440
Qy	1441	CAACAACTCGGAATTCAGATCTTAAACAAAGATATGATTCGAGGAGCCCGGCAAAAT	1500
Db	1441	CAACAACTCGGAATTCAGATCTTAAACAAAGATATGATTCGAGGAGCCCGGCAAAAT	1500
Qy	1501	TTTTCAGCTTATGGGTAGGATGCAATCGAGCTGACAATAACCATGATTTTACAGTTACAT	1560
Db	1501	TTTTCAGCTTATGGGTAGGATGCAATCGAGCTGACAATAACCATGATTTTACAGTTACAT	1560
Qy	1561	ACAACTTTTAAATGTTTATACGAGGAGCTCACTGTTCTAGTGTGAGGGATATCGGCTT	1620
Db	1561	ACAACTTTTAAATGTTTATACGAGGAGCTCACTGTTCTAGTGTGAGGGATATCGGCTT	1620
Qy	1621	CTTAGTATTTGGATGAATCATCAACAACTATTTTAAAGTGTTCAGAACATAAAGAG	1680
Db	1621	CTTAGTATTTGGATGAATCATCAACAACTATTTTAAAGTGTTCAGAACATAAAGAG	1680
Qy	1681	GAAATGAGCCCTGTAAAGACTATATATGAGGACCATCAATATCGCGCCCGCAATTC	1737
Db	1681	GAAATGAGCCCTGTAAAGACTATATATGAGGACCATCAATATCGCGCCCGCAATTC	1737

Dn	1021	AGAGGTGCAACAATTTATTTGCGCCAGAGAGTTTGCAGAACATATATATTTTGGTGAGCATGGT	1080
Qy	1081	TCTAGTTTGGGCGAGTTTTCAGCAGATATCTTGAGGCAATTAACCTAAATGATGTCCAG	1140
Dn	1081	TCTAGTTTGGGCGAGTTTTCAGCAGATATCTTGAGGCAATTAACCTAAATGATGTCCAG	1140
Qy	1141	GTTCATTTGGAAGTCAATGGACCTTAGTTACCTTTGGAGGACCAATTAACGTGAAACACTTT	1200
Dn	1141	GTTCATTTGGAAGTCAATGGACCTTAGTTACCTTTGGAGGACCAATTAACGTGAAACACTTT	1200
Qy	1201	GGTGACTTGGTTAAAAAGGCTTAAGCCATCCATGGAGCTGATGCTGCTTTGAAAGCAATTT	1260
Dn	1201	GGTGACTTGGTTAAAAAGGCTTAAGCCATCCATGGAGCTGATGCTGCTTTGAAAGCAATTT	1260
Qy	1261	AACATAGATGGTATGTGCGTATTCAGTACAGAGATCAACTAGACTTTTGAAGATATCGCA	1320
Dn	1261	AACATAGATGGTATGTGCGTATTCAGTACAGAGATCAACTAGACTTTTGAAGATATCGCA	1320
Qy	1321	CGGCAATTTGGCATTTTGAAGATGGAAGATGGTGTACCCAGCTGCAGCATATAAAGGA	1380
Dn	1321	CGGCAATTTGGCATTTTGAAGATGGAAGATGGTGTACCCAGCTGCAGCATATAAAGGA	1380
Qy	1381	ATAGTAGTTTCCGGTACCAACCTCCAGAGCTGATTTCTTGTGGCCATGATTCGCTT	1440
Dn	1381	ATAGTAGTTTCCGGTACCAACCTCCAGAGCTGATTTCTTGTGGCCATGATTCGCTT	1440
Qy	1441	CAACAACCTGGAATGGAAGATCTTAACAAAGATATGATTCAGAGGCGCGGGCAAAATTT	1500
Dn	1441	CAACAACCTGGAATGGAAGATCTTAACAAAGATATGATTCAGAGGCGCGGGCAAAATTT	1500
Qy	1501	TTTGACTTATTTGGGTAGGATGCATCGAGCTGACACTAAACCATCATTTTACCAGTTACAT	1560
Dn	1501	TTTGACTTATTTGGGTAGGATGCATCGAGCTGACACTAAACCATCATTTTACCAGTTACAT	1560
Qy	1561	ACAAAGTTTAAATGTTATACGGAGGAGCTCACTGTTCTAGTGTGTAAGGGATATCGGCTT	1620
Dn	1561	ACAAAGTTTAAATGTTATACGGAGGAGCTCACTGTTCTAGTGTGTAAGGGATATCGGCTT	1620
Qy	1621	CTTAGTATGGAATGAAATCAACACACACCTATTATTTTAAAGTGTTCAGAACATAAAGAG	1680
Dn	1621	CTTAGTATGGAATGAAATCAACACACACCTATTATTTTAAAGTGTTCAGAACATAAAGAG	1680
Qy	1681	GAAATGAGCCCTTAAGACTATACATGGGACCATCATATTCGGGCGCGGAAATTC	1737
Dn	1681	GAAATGAGCCCTTAAGACTATACATGGGACCATCATATTCGGGCGCGGAAATTC	1737

RESULT 3
AR432619 LOCUS AR432619 1737 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 3 from patent US 6653459.
ACCESSION AR432619
VERSION AR432619.1 GI:40195152
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1737)
AUTHORS Von Schaewen,A.
TITLE Plant Gnti sequences and the use thereof for the production of
plants having reduced or lacking N-acetyl glucosaminyl transferase
I(GntI) activity
JOURNAL Patent: US 6653459-A 3 25-NOV-2003;
FEATURES Location/Qualifiers
source 1..1737
/organism="unknown"
/mol_type="genomic DNA"

= parent of instant app.

Query Match 100.0%; Score 1737; DB 6; Length 1737;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1737; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GAATTTCGGCGCCGCAATTTGACTTCCTAACTGAAACAGGCAAGTAAATCCAGCATGA	60
Dn	1	GAATTTCGGCGCCGCAATTTGACTTCCTAACTGAAACAGGCAAGTAAATCCAGCATGA	60
Qy	61	AACACTCATAACTGAACACACTGAGAGACTATTCCTCTAAAGCCTTCAATCGAATTT	120
Dn	61	AACACTCATAACTGAACACACTGAGAGACTATTCCTCTAAAGCCTTCAATCGAATTT	120
Qy	121	CGCAGATGAGAGGGAACAAGTTTTCGTGTGATTTCCGGTACCTCTCATCTTGGCTGCT	180
Dn	121	CGCAGATGAGAGGGAACAAGTTTTCGTGTGATTTCCGGTACCTCTCATCTTGGCTGCT	180
Qy	181	GTGCGCTTCATCTACACACAGATGCGGCTTTTTCGACACAGTCAGATATGAGATCGC	240
Dn	181	GTGCGCTTCATCTACACACAGATGCGGCTTTTTCGACACAGTCAGATATGAGATCGC	240
Qy	241	CTTGCTGCTGCAATTTGAAGCAGAAAAATCAATTTGTAAGCCAGACAGATTGCTTATTGAC	300
Dn	241	CTTGCTGCTGCAATTTGAAGCAGAAAAATCAATTTGTAAGCCAGACAGATTGCTTATTGAC	300
Qy	301	CAGATTAGCTGAGCAAGAAATAGTTGCTCTTTGAAAGAAACAAATGAAGCGTCAGGAC	360
Dn	301	CAGATTAGCTGAGCAAGAAATAGTTGCTCTTTGAAAGAAACAAATGAAGCGTCAGGAC	360
Qy	361	CAGGAGTCCGCAATTTAAGGGCTCTTGTTCAGGATCTTGAAGTGAAGGCAATAAAG	420
Dn	361	CAGGAGTCCGCAATTTAAGGGCTCTTGTTCAGGATCTTGAAGTGAAGGCAATAAAG	420
Qy	421	TTGATCGGAAATGTACAGATGCCAGTGGCTGTGTAGTTGTGTATGGCTTCAATTCGGGCT	480
Dn	421	TTGATCGGAAATGTACAGATGCCAGTGGCTGTGTAGTTGTGTATGGCTTCAATTCGGGCT	480
Qy	481	GATTACCTGGAAGAAGCTATTAAATCCATCTTAAATACCAATATCTGTGTGGTCAAAA	540
Dn	481	GATTACCTGGAAGAAGCTATTAAATCCATCTTAAATACCAATATCTGTGTGGTCAAAA	540
Qy	541	TATCTCTTTTTCATATCCAGGATGCATCATCTGATGTCCAGGAGCTTGTCTTCAGC	600
Dn	541	TATCTCTTTTTCATATCCAGGATGCATCATCTGATGTCCAGGAGCTTGTCTTCAGC	600
Qy	601	TATGATCAGCTGAGTATATGACGACCTTTGGAATTTTGAACCTGTGCATATCTGAAAGACCA	660
Dn	601	TATGATCAGCTGAGTATATGACGACCTTTGGAATTTTGAACCTGTGCATATCTGAAAGACCA	660
Qy	661	GGGAGCTGATTCATACATAAATTTGCAAGTGGCAAGTGGCAAGTGGATCGCTG	720
Dn	661	GGGAGCTGATTCATACATAAATTTGCAAGTGGCAAGTGGCAAGTGGATCGCTG	720
Qy	721	TTTTCAAGCATAAATTTAGCCCTGTATCATACTAGAGATCATATGGAATTTGCCCT	780
Dn	721	TTTTCAAGCATAAATTTAGCCCTGTATCATACTAGAGATCATATGGAATTTGCCCT	780
Qy	781	GATTTTTCGACCTTTTTCAGGCTGGAGCTACTCTTCTTCAGCAGAGCAAGTCGATTATG	840
Dn	781	GATTTTTCGACCTTTTTCAGGCTGGAGCTACTCTTCTTCAGCAGAGCAAGTCGATTATG	840
Qy	841	GCTATTTCTCTTGGAAATGACAAATGAGCAATGCGATTTTGTCCAAGATTCCTTATGCTCTT	900
Dn	841	GCTATTTCTCTTGGAAATGACAAATGAGCAATGCGATTTTGTCCAAGATTCCTTATGCTCTT	900
Qy	901	TACCGCTCAGATTTTTCGCGCTTGGATGATGCTTTCAAAATCTACTTGGGACGAA	960
Dn	901	TACCGCTCAGATTTTTCGCGCTTGGATGATGCTTTCAAAATCTACTTGGGACGAA	960
Qy	961	TTATCTCCAAAGTGGCCAAAGGCTTACTGGGACGCTGGCTTAAGACTCAAGAGAAATCAC	1020
Dn	961	TTATCTCCAAAGTGGCCAAAGGCTTACTGGGACGCTGGCTTAAGACTCAAGAGAAATCAC	1020
Qy	1021	AGAGGTGCAACAATTTATTCGCCAGAGGTTTTCAGAACATATATTTTGGTGAGCATGGT	1080
Dn	1021	AGAGGTGCAACAATTTATTCGCCAGAGGTTTTCAGAACATATATTTTGGTGAGCATGGT	1080
Qy	1081	TCTAGTTTGGGCGAGTTTTCAGCAGATATCTTGAGGCAATTAACCTAAATGATGTCCAG	1140

615 GTATATGACGACCTTGGATTTTGAACCTGTGTCATCTGAAGACCCAGGGAGCTGATTGC 674
Db GTATATGACGACCTTGGATTTTGAACCTGTGTCATCTGAAGACCCAGGGAGCTGATTGC 660
675 ATACTACAAAATTCGACCTCATTAACAAGTGGGCAATGGATCAGCTGTTTTCACAAGCATAA 734
Db ATACTACAAAATTCGACCTCATTAACAAGTGGGCAATGGATCAGCTGTTTTCACAAGCATAA 720
735 TTTTAGCGGTGTTATCATPACTAGAGATGATPAGAAATTCGCCCTGATTTTTTTGACTT 794
Db TTTTAGCGGTGTTATCATPACTAGAGATGATPAGAAATTCGCCCTGATTTTTTTGACTT 780
795 TTTTAGCGGTGACTACTCTCTTGACAGAGACAAGTCGATATGCTATTTCTTCTT 854
Db TTTTAGCGGTGACTACTCTCTTGACAGAGACAAGTCGATATGCTATTTCTTCTT 840
855 GAATGACAATGACAAATGCAATTTGTGTCAGAGATCCTTATGCTCTTTACCGCTCAGATTT 914
Db GAATGACAATGACAAATGCAATTTGTGTCAGAGATCCTTATGCTCTTTACCGCTCAGATTT 900
915 TTTTCCCGCTTGGATGATGCTTTCAAATCTACTTTGGAGCAATATCTCCAAAGTG 974
Db TTTTCCCGCTTGGATGATGCTTTCAAATCTACTTTGGAGCAATATCTCCAAAGTG 960
975 GCCAAGGCTTACTGGGACGACTGGCTAAGACTCAAGAGNATCACAGAGTGCACAAAT 1034
Db GCCAAGGCTTACTGGGACGACTGGCTAAGACTCAAGAGNATCACAGAGTGCACAAAT 1020
1035 TATTGCGCCAGAGTTTGCAGAACATATAATTTTGGTGAGCATGTTCTAGTTTGGGGCA 1094
Db TATTGCGCCAGAGTTTGCAGAACATATAATTTTGGTGAGCATGTTCTAGTTTGGGGCA 1080
1095 GTTTTTCAGCAGTATCTTGAGCCAAATPAAACTAAATGATGTCAGGTGATTTGGAAGTC 1154
Db GTTTTTCAGCAGTATCTTGAGCCAAATPAAACTAAATGATGTCAGGTGATTTGGAAGTC 1140
1155 AATGACCTTAGTTACCTTTTGGAGGACAAATGATGTAACACACACACACACACACAC 1214
Db AATGACCTTAGTTACCTTTTGGAGGACAAATGATGTAACACACACACACACACACAC 1200
1215 AAAGGCTAAGCCCATCCATGAGCTGATGCTGCTTGAAGCAATTAACATAGATGTTGA 1274
Db AAAGGCTAAGCCCATCCATGAGCTGATGCTGCTTGAAGCAATTAACATAGATGTTGA 1260
1275 TGTGCGTATTTCAGTACAGAGATCAACTAGACTTTGAAAAATATCGCACGCAATTTGGCAT 1334
Db TGTGCGTATTTCAGTACAGAGATCAACTAGACTTTGAAAAATATCGCACGCAATTTGGCAT 1320
1335 TTTTGAAGATGGAAGGATGTTGATCCACGTCGAGCATATAAGGAATAGTAGTTTTCG 1394
Db TTTTGAAGATGGAAGGATGTTGATCCACGTCGAGCATATAAGGAATAGTAGTTTTCG 1380
1395 GTACCAACGCTCCAGACGCTGATTTCTTTTGGCCATGATTCGCTTCAACAACTCGCAAT 1454
Db GTACCAACGCTCCAGACGCTGATTTCTTTTGGCCATGATTCGCTTCAACAACTCGCAAT 1440
1455 TGAAGATACCTTAAACAAGATATGATTCAGGAGCCCGGCAAAATTTTTCATTTATGGG 1514
Db TGAAGATACCTTAAACAAGATATGATTCAGGAGCCCGGCAAAATTTTTCATTTATGGG 1500
1515 TAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1574
Db TAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
1575 TTATACGAGGAGCTCAGCTGTTCTAGTGTTCAGAGGATATCGGCTCTTCTAGTATGGATG 1634
Db TTATACGAGGAGCTCAGCTGTTCTAGTGTTCAGAGGATATCGGCTCTTCTAGTATGGATG 1620
1635 AATCATCAACACCACTTATTTTAAAGTGTTCAGAACATAAAGAGGAAATGTTAGCCCTG 1694
Db AATCATCAACACCACTTATTTTAAAGTGTTCAGAACATAAAGAGGAAATGTTAGCCCTG 1680

1695 TAAAGACTATACATGGGACCATCAAT 1722
Db TAAAGACTATACATGGGACCATCAAT 1708
RESULT 5
NTY16832
LOCUS
DEFINITION
NTY16832 2031 bp mRNA linear PLN 18-APR-2005
Nicotiana tabacum mRNA for
beta-1,2-N-acetylglucosaminyltransferase, complete CDS.
ACCESSION
VERSION
Y16832.1 GI:5764048
KEYWORDS
beta-1,2-N-acetylglucosaminyltransferase.
SOURCE
Nicotiana tabacum (common tobacco)
ORGANISM
Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE
1 Strasser, R., Mucha, J., Schwihla, H., Altmann, F., Glessl, J. and
Steinkellner, H.
Molecular cloning and characterization of cDNA coding for beta1,
2N-acetylglucosaminyltransferase I (GlcNAc-TI) from Nicotiana
tabacum
Glycobiology 9 (8), 779-785 (1999)
REFERENCE
2 (bases 1 to 2031)
Steinkellner, H.
Direct Submission
Submitted (06-MAR-1998) H. Steinkellner, Zentrum fuer Angewandte
Genetik, BOKU-Wien, Muthgasse 18, 1190 Wien, AUSTRIA
FEATURES
source
/organism="Nicotiana tabacum"
/mol_type="mRNA"
/cultivar="Xanthi"
/db_xref="taxon:4097"
1..76
77..1417
/EC_number="2.4.1.101"
/codon_start=1
/product="alpha-1,3-mannosyl-glycoprotein
beta-1,2-N-acetylglucosaminyltransferase"
/protein_id="CA553347.1"
/db_xref="GI:5764048"
/db_xref="GOA:O9ST97"
/db_xref="InterPro:IPR001173"
/db_xref="InterPro:IPR004139"
/db_xref="UniProt/TREMBL:O9ST97"
/translation="MRGYKFCDFRYLLILAAVAFIYIQWRLPATSEYADRLAAALIE
AENHCTSQRLIIDIQISQGGRIVALDEQMKQDQRCQRLRALVQDLSEKIGKILGN
VQMPVAHVVMACNRADYLEIKTSILKYQISVAPKPLIFISQDGHDPVILDKLMSYD
QUTYMHLDPEPVHTPERGELIAYYKIHRYKWDQLFYKHNFSRVILLEDMEIAP
DFDFPEAGATLLDRKSIATSSWNDQMGQFVQYALYRSDFFGLGWLMSKSTW
DELSKPWKPYAWDDWLRLKENHRGROFIRPEVCSYNGEHSGLGQFPKQYLEPIKL
NDVQDWKSNLSDLYLLEDNYVHFGLVKAPDKIHGDAVLKAPNIDGVDRIQYRDOL
NDVQDWKSNLSDLYLLEDNYVHFGLVKAPDKIHGDAVLKAPNIDGVDRIQYRDOL
DIADIARQGIPEEWMKGVPRAAIKGIVVFRYQTSRRVFLVGDPSLQQUGNEDI"
1418..2031
3' UTR
ORIGIN
Query Match 89.9%; Score 1561.8; DB 15; Length 2031;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 1627; Conservative 0; Mismatches 47; Indels 7; Gaps 3;
51 CCAGCGATGAACACTCAATACTGAACACTGAGAGACTATTCGCTTTCTCCTAAAGCCTT 110
Db CCAGCGATGAACACTCAATACTGAACACTGAGAGACTATTCGCTTTCTCCTAAAGCCTT 60
111 CAATCGAATTCGACGATGAGAGGAAACAAGTTTTCGTGTGATTTCCGGTACCTCCTCAT 170
Db CAATCGAATTCGACGATGAGAGGAAACAAGTTTTCGTGTGATTTCCGGTACCTCCTCAT 120
171 CTTGGCTGCTGCTGCCTTCATCTACACACAGATCGCGCTTTTTCGCGACACAGTACAGATA 230

121 CTTGGCTGCTGCGCTTCATCTACATACAGATCGGCTTTTCGGACACAGTCAGAATA 180
1231 TGCAGATCGCCTTGCCTGCTGCAATTTGAAGCAGAAAAATCATTTGTACAAGCCAGACCAAT 290
1261 TGCAGATCGCCTTGCCTGCTGCAATTTGAAGCAGAAAAATCATTTGTACAAGCCAGACCAAT 240
1291 GCTTATTGACAGATTTAGCTGCGCAAGCAAGAAATAGTTGCTCTTGAAGAAACAAATGAA 350
1321 GCTTATTGACAGATTTAGCTGCGCAAGCAAGAAATAGTTGCTCTTGAAGAAACAAATGAA 300
1351 GCCTCAGGACAGGAGTCCGACAAATTAAGGGCTCTTGTTCAGGATCTTGAAGTAAGGG 410
1381 GCCTCAGGACAGGAGTCCGACAAATTAAGGGCTCTTGTTCAGGATCTTGAAGTAAGGG 360
1411 CATAAAAAGTTGATCGGAATGTACAGATGCGAGTGGCTGCTGCTAGTTGTATGGCTTG 470
1441 CATAAAAAGTTGATCGGAATGTACAGATGCGAGTGGCTGCTGCTAGTTGTATGGCTTG 420
1471 CAATCGGCTGATTTACCTGGAAGAGACTATTAAATCCCATCTTAAATACCAAAATATCTGT 530
1501 CAATCGGCTGATTTACCTGGAAGAGACTATTAAATCCCATCTTAAATACCAAAATATCTGT 480
1531 TGCGTCAAAATATCTCTTTTCAATCCAGGATGGATCACATCCTGATGTCAGGAAGCT 590
1561 TGCGGCAAAATATCTCTTTTCAATCCAGGATGGATCACATCCTGATGTCAGGAAGCT 540
1591 TGCTTTGAGCTATGATCAGCTGACGTAATGCAATGCAAAATGCAAAATGCAAAATGCAAA 650
1621 TGCTTTGAGCTATGATCAGCTGACGTAATGCAATGCAAAATGCAAAATGCAAAATGCAAA 600
1651 TGAAAGACAGGGAGCTGATGTCATCTACAAATGCAAAATGCAAAATGCAAAATGCAAAAT 710
1681 TGAAAGACAGGGAGCTGATGTCATCTACAAATGCAAAATGCAAAATGCAAAATGCAAAAT 660
1711 GGATCAGCTGTTTTCAAGACATAATTTTAGCGCTGTTATCATCTACATGAGATGATATGGA 770
1741 GGATCAGCTGTTTTCAAGACATAATTTTAGCGCTGTTATCATCTACATGAGATGATATGGA 720
1771 AATGCCCTGATTTTTTGACTTTTTTGAGGCTGGAGCTACTCTTCTTGACAGAGCAA 830
1801 AATGCCCTGATTTTTTGACTTTTTTGAGGCTGGAGCTACTCTTCTTGACAGAGCAA 780
1831 GTCGATTTAGCTATTTCTTTGCAATGCAATGCAAAATGCAAAATGCAAAATGCAAAATGCA 890
1861 GTCGATTTAGCTATTTCTTTGCAATGCAATGCAAAATGCAAAATGCAAAATGCAAAATGCA 840
1891 TTATGCTCTTTACCGCTCAGATTTTTTCCCGCTCTTGGATGGATGCTTTCAAAATCTAC 950
1921 TTATGCTCTTTACCGCTCAGATTTTTTCCCGCTCTTGGATGGATGCTTTCAAAATCTAC 900
1951 TTGGGACGAATTTATCTCAAGTGGCCAAAGGCTTACTGGGACGACTGGCTAAGACTCAA 1010
1981 TTGGGACGAATTTATCTCAAGTGGCCAAAGGCTTACTGGGACGACTGGCTAAGACTCAA 960
2011 AGAAGATCAGAGGTCGCAATTTATTCGCCAGAGTTTTCGAGACATATAATTTGG 1070
2041 AGAAGATCAGAGGTCGCAATTTATTCGCCAGAGTTTTCGAGACATATAATTTGG 1020
2071 TCAGCATGCTTCTAGTTTGGGCGAGTTTTTCAAGCAGATCTTGGCCAAATTTAAATCAA 1130
2101 TCAGCATGCTTCTAGTTTGGGCGAGTTTTTCAAGCAGATCTTGGCCAAATTTAACTAA 1080
2131 TGATGTCAGAGTTGATTTGAAGTCAATGGACCTTATGTTACCTTTTGGAGGACAAATTA 1190
2161 TGATGTCAGAGTTGATTTGAAGTCAATGGACCTTATGTTACCTTTTGGAGGACAAATTA 1140
2191 GAAACACTTTTGGTGAATTTGTTTAAAGGCTAAGCCCATCCATGGAGCTGATGCTGCTT 1250
2221 GAAACACTTTTGGTGAATTTGTTTAAAGGCTAAGCCCATCCATGGAGCTGATGCTGCTT 1200
2251 GAAAGCATTTAATAGATGCTGATGCTGATTTAGTACAGATCAAGATCAACTAGATTTCA 1310

1201 GAAAGCATTTTAAATAGATGGTGAATGTCGCTATTTAGTACAGAGATCAACTAGACTTTGA 1260
1311 AAATATCGCACGGCAATTTTGGCAATTTTGAAGAAATGGAAGATGGTGTACACGTCAGC 1370
1261 AGATATCGCACGGCAATTTTGGCAATTTTGAAGAAATGGAAGATGGTGTACACGTCAGC 1320
1371 ATATAAGGAATAGTAGTTTCCGGTACCAAAAGTCCAGACGTCATTTCTTGTGGCCA 1430
1321 ATATAAGGAATAGTAGTTTCCGGTACCAAAAGTCCAGACGTCATTTCTTGTGGCCC 1380
1431 TGATTCGCTTCAACCACTCGGAATTTGAAGATCTTTAAACAAGATATGATTT---GGAGCCC 1490
1381 TGATTCGCTTCAACCACTCGGAATTTGAAGATCTTTAAACAAGATATGATTT---GGAGCCC 1437
1491 GGGCAAAATTTTTCGCTTATTTGGTAGGATCGATCGAGCTGACACTAAACCATGATTTTA 1550
1438 GGACAAAGATTTAGACTTATTTGGTAGGATCGATCGAGCTGACACCAACCATGATTTTA 1497
1551 CCAGTTACATACAAACGTTTAA---TGTTATACGGAGGAGCTCAGCTGTTCTAGTGTGAAG 1609
1498 CCAGTTACATACAAACGTTTAA---TGTTATACGGAGGAGCTCAGCTGTTCTAGTGTGAAG 1557
1610 GATATCGCTTCTTAGTATTTGGATGATCATCAACCAACCTATTATT---TTAAGTCTT 1666
1558 GATATCGCTTCTTAGTATTTGGATGATCATCAACCAACCTATTATT---TTAAGTCTT 1617
1667 CAGAACATAAGAGGAAATGTAGCCCTGTAAAGACTATACATGGGACCATCATATCCG 1726
1618 CCGAACATAAGAGGAAATGTAGCCCTGTAAAGACTATACATGGGACCATCATATCCG 1677
1727 G 1727
1678 G 1678

RESULT 6
NTA249882
LOCUS NTA249882 1836 bp mRNA linear PLN 15-APR-2005
DEFINITION Nicotiana tabacum mRNA for N-acetylglucosaminyltransferase I (GntI gene), clone A4.
ACCESSION AJ249882.1 GI:18076145
VERSION GntI gene; N-acetylglucosaminyltransferase I.
KEYWORDS Nicotiana tabacum (common tobacco)
SOURCE Nicotiana tabacum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE 1 Wenderoth, I., Tjaden, J. and von Schaewen, A.
Isolation and characterization of different plant N-acetylglucosaminyltransferase I (GntI) cDNA sequences: Functional analyses in the Arabidopsis cgl mutant, and in potato and tobacco antisense plants
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1836)
AUTHORS von Schaewen, A.
TITLE Direct Submission
JOURNAL Submitted (28-SEP-1999) von Schaewen A., Pflanzenphysiologie, Universitaet Osnabrueck, FB5, Barbarastr. 11, D-49076 Osnabrueck, GERMANY
FEATURES
source Location/Qualifiers
1.1836 /organism="Nicotiana tabacum"
/mol_type="mRNA"
/cultivar="Samsun NN"
/db_xref="taxon:4097"
/clone="A4"
/tissue_type="leaf mesophyll"
/note="Sink"
1.1836 /gene="GntI"
27.1367 CDS

SOURCE Nicotiana benthamiana
 ORGANISM Nicotiana benthamiana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Nicotiana.

REFERENCE 1
 AUTHORS Strasser, R., Glossl, J. and Steinkellner, H.
 TITLE Less than 2% of GntI activity does not alter N-glycosylation in Nicotiana benthamiana
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1341)
 AUTHORS Steinkellner, H.
 TITLE Direct Submission
 JOURNAL Submitted (20-SEP-2000) Steinkellner H., University of Agricultural Sciences-Vienna, Center of Applied Genetics, Muthgasse 18, 1190 Vienna, AUSTRIA

FEATURES
 source
 1. .1341
 /organism="Nicotiana benthamiana"
 /mol_type="mRNA"
 /db_xref="taxon:4100"
 1. .1341
 /gene="gntI"
 76. .1341
 /gene="gntI"
 /codon_start=1
 /product="N-acetylglucosaminyltransferase I"
 /protein_id="CAC82507.1"
 /db_xref="GI:27527069"
 /db_xref="GOA:Q8GUAS"
 /db_xref="InterPro:IPR004139"
 /db_xref="UniProt/TREMBL:Q8GUAS"
 /translation="MRLPATQSEVADRLRAAIEAENHCTSOIRLLIDHISQOGRIVA LEEQKQDQRCRLVQLQESKGIKGLGNVQMPVAVVWACSRADYLEKTIK MLKYQVSVAQYFISQDGHDPVKLALSYDQTYMHQDLPFVHTRFEGELIAY KIARHWALDQLFYKHFNSRVIILEDMEIAPDFDFEAGATLLDRDKSINAISS NDNGQMOQPYALYRSDLPFGLGMLSKSTWDELSPKPKAYMDWLRLKINHRGR QFIPVCRTYNFGEHSSLGQFPKQVLEPIKLANDVOVDWKSMDLSYLLEONYVKHFG NLVKAKEPIHGADAVLKAFNIDGVDVRIQYRQLEFEDIARQGIFFSEKDGVPRAAYK GIVFWYOTSRVRFLVGHDSLQQLGIEDT"

ORIGIN
 Query Match 74.7%; Score 1297.8; DB 15; Length 1341;
 Best Local Similarity 98.0%; Pred. No. 7,7e-307;
 Matches 1314; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 127 ATGAGAGGGAACAAGTTTGTCTGTGATTTCGGTACCTCCTCATCTTGGCTGCTGC 186
 DB 1 ATGAGAGGGAACAAGTTTGTCTGTGATTTCGGTACCTCCTCATTTTGGCTGCTGC 60
 QY 187 TTCATCTACACAGATCGGCTTTTGGGACACAGTCAGATATGCAGATCGCCTTGT 246
 DB 61 TTCATCTACATACAGATCGGCTTTTGGGACACAGTCAGATATGCAGATCGCCTTGT 120
 QY 247 GCTGCAATTGAGCAGAAATCATTTGTAAGCAGACAGATTCGTTATTGACAGATT 306
 DB 121 GCTGCAATTGAGCAGAAATCATTTGTAAGCAGACAGATTCGTTATTGATCATATT 180
 QY 307 AGCCTGCAGCAAGGAAGATAGTTGCTCTTGAAGAACAAATGAAGCGTCAGGACGAG 366
 DB 181 AGCCAGCAGCAAGGAAGATAGTTGCTCTTGAAGAACAAATGAAGCGCCAGGACGAG 240
 QY 367 TCCGACAAATTAAGGCTTCTGTTCCAGATCTTTGAAGTAAGGGCATATAAAGTTGATC 426
 DB 241 TCCGACAAATTAAGGCTTCTGTTCCAGATCTTTGAAGTAAGGGCATATAAAGTTGATC 300
 QY 427 GGAATGTACAGATCGCAGTGCCTGTAGTTGTTATGCTTGCATCGGCTGATTAC 486
 DB 301 GGAATGTACAGATCGCAGTGCCTGTAGTTGTTATGCTTGCATCGGCTGACTAC 360
 QY 487 CTGGAAGAGACTATTAAATCCATCTTAAATAACCAATATCTGTTGCGTCAAAATATCT 546
 DB 361 CTGGAAGAGACTATTAAATCCATCTTAAATAACCAATATCTGTTGCGTCAAAATATCT 420

QY 547 CTTTTCATATCCAGAGTGGATCACATCCTGATGTTCAGGAAGCTTGTCTTTCAGCTATGAT 606
 DB 421 CTTTTCATATCCAGAGTGGATCACATCCTGATGTTCAGGAAGCTTGTCTTTCAGCTATGAT 480
 QY 607 CAGTGCATATATGACAGCACTTGGATTTTGAACCTGTGATCTACTGAAAGACCGGGAG 666
 DB 481 CAGTGCATATATGACAGCACTTGGATTTTGAACCTGTGATCTACTGAAAGACCGGGAG 540
 QY 667 CTGATTGCATATACAAAATTGCAAGTGGCATTTACAAAGTGGCATTTGGATTCAGCTCTTTTAC 726
 DB 541 CTGATTGCATATACAAAATTGCAAGTGGCATTTACAAAGTGGCATTTGGATTCAGCTCTTTTAC 600
 QY 727 AAGCATAAATTTAGCCGCTGTTATCATATACTAGAAGATGATATGGAATTTGCCCTTGATTTT 786
 DB 601 AAGCATAAATTTAGCCGCTGTTATCATATACTAGAAGATGATATGGAATTTGCCCTTGATTTT 660
 QY 787 TTTGACATTTTGGAGGCTGGAGCTACTCTTCTTGACAGAGCAAGTCGATTATGCTATT 846
 DB 661 TTTGACATTTTGGAGGCTGGAGCTACTCTTCTTGACAGAGCAAGTCGATTATGCTATT 720
 QY 847 TCTTCTTGAATGACAAATGCAAAATGCAGTTTGTCCAAGATCCTTATGCTCTTTTACCGC 906
 DB 721 TCTTCTTGAATGACAAATGCAAAATGCAGTTTGTCCAAGATCCTTATGCTCTTTTACCGC 780
 QY 907 TCAGATTTTTCCTCGGTCTTGGATGGATGCTTTCAAAATCTACTTGGGACGAATTTATCT 966
 DB 781 TCAGATTTTTCCTCGGTCTTGGATGGATGCTTTCAAAATCTACTTGGGACGAATTTATCT 840
 QY 967 CCAAAGTGGCAAGGCTTACTTGGGACGATGGCTTAAGACTCAAGAGAAATCACAGAGGT 1026
 DB 841 CCAAAGTGGCAAGGCTTACTTGGGACGATGGCTTAAGACTCAAGAGAAATCACAGAGGT 900
 QY 1027 CGACAAATTTATTCGCCCAGAAAGTTTGCAGAACATATAATTTTGTGTGAGCATGGTTCTAGT 1086
 DB 901 CGACAAATTTATTCGCCCAGAAAGTTTGCAGAACATATAATTTTGTGTGAGCATGGTTCTAGT 960
 QY 1087 TTGGGGCAGTTTTTCAAGCAGTATCTTGGAGCAATTTAAACTTAAATGATGTCCAGGTTGAT 1146
 DB 961 TTGGGGCAGTTTTTCAAGCAGTATCTTGGAGCAATTTAAACTTAAATGATGTCCAGGTTGAT 1020
 QY 1147 TGGAGTCAATGGAGCTTAGTTTACTCTTTGGAGGACAAATTCGTGAAACACTTTGGTGAC 1206
 DB 1021 TGGAGTCAATGGAGCTTAGTTTACTCTTTGGAGGACAAATTCGTGAAACACTTTGGTGAC 1080
 QY 1207 TTGTTTAAAGAGGCTAAGCCCATCCATCGAGCTGATGCTGCTTGAAGAGCAATTTAAACATA 1266
 DB 1081 TTGTTTAAAGAGGCTAAGCCCATCCATCGAGCTGATGCTGCTTGAAGAGCAATTTAAACATA 1140
 QY 1267 GATGCTGATGTGCGTATTTCAGTACAGAGATCAACTAGACTCTTTGAAATAATTCGACGGCAA 1326
 DB 1141 GATGCTGATGTGCGTATTTCAGTACAGAGATCAACTAGACTCTTTGAAATAATTCGACGGCAA 1200
 QY 1327 TTTGGCATTTTGAAGAAATGGAAGATGGTGTACCACTGAGCATATAAAGGAATAGTA 1386
 DB 1201 TTTGGCATTTTGAAGAAATGGAAGATGGTGTACCACTGAGCATATAAAGGAATAGTA 1260
 QY 1387 GTTTTCCGTTACCAACGTCAGAGCTGATTTTCCCTTGTGGCATGATTTCCGTTCAACAA 1446
 DB 1261 GTTTTCCGTTACCAACGTCAGAGCTGATTTTCCCTTGTGGCATGATTTCCGTTCAACAG 1320
 QY 1447 CTCGGAATTGAAGATACCTTAA 1467
 DB 1321 CTCGGAATTGAAGATACATAA 1341

RESULT 8
 STU249878
 LOCUS Solanum tuberosum mRNA for N-acetylglucosaminyltransferase I (GntI gene), clone A1.
 DEFINITION Solanum tuberosum mRNA for N-acetylglucosaminyltransferase I (GntI gene), clone A1.
 ACCESSION AJ249878
 VERSION AJ249878.1 GI:18076139

Db 1358 ACTTGGAGTTGAAGTACTTAGCAGAGATATGATT---GGAGCCTGAGCAACAATTTAGA 1414
Qy 1506 CTTATTGGGTAGGATGCATCGAGCTGACACTAAACCATGATTTT-----ACCAGTTAC 1558
Db 1415 CTTATTGGTGGTACATTTTGAAGAGCTGACACGAAAGATATGACTACCAAGTAGCTAC 1474
Qy 1559 ATACAACGTTTTAATGTTATACGAGGAGCTCACTGTTCTAGTGTGAAGGATATCGGC 1618
Db 1475 ATGCAACATTTAATGTTAATGGAAGGACCACTGCT-----T 1513
Qy 1619 TTCTTAGTATTGGATGATATCAACAACTATTATTATTTAAGTGTTCAGAACATAAG 1678
Db 1514 ATTGTTGGAATGATGATATCATCCACATCTATTA--TTCAAAGTTTACAAACATAAG 1571
Qy 1679 AGGAATGTAGCCCTGTAAAGACTA 1703
Db 1572 AGGAATGTTCCTCTATAAAACAA 1596

RESULT 9

A95053 LOCUS 1669 bp DNA linear PAT 26-JAN-2000
DEFINITION Sequence 1 from Patent WO929879.
ACCESSION A95053
VERSION A95053.1 GI:6779205

KEYWORDS Solanum tuberosum (potato)

SOURCE Solanum tuberosum

ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 1669)

AUTHORS Von,S.A.

TITLE VEGETABLE GntI SEQUENCES AND THE USE THEREOF TO OBTAIN PLANTS WITH
A REDUCED OR LACK OF N-ACETYLGLUCOSAMINYLTRANSFERASE I (GntI)

ACTIVITY

JOURNAL Patent: WO 929879-A 1 17-JUN-1999;

AUTHORS VON SCHAEWEN ANTJE (DE)

FEATURES Location/Qualifiers

source

1..1669

/organism="Solanum tuberosum"

/mol_type="unassigned DNA"

/cultivar="Desiree"

/db_xref="taxon:4113"

/cell_type="BLATTZELLEN"

/tissue_type="MESOPHYLL"

/clone_lib="LAMBDA ZAP II (ECORI)"

/dev_stage="M: 'SINK' ORGAN"

15..52

53..1393

/gene="CGL"

/EC_number="2.4.1.101"

53..1393

/gene="CGL"

/function="INITIERT KOMPLEXE N-GLYKANE AUF SEKRET."

GLYKOPROTEINEN"

/standard_name="GNTI"

/note="ERSTE GNTI-SEQUENZ AUS KARTOFFEL (UNPUBLIZIERT)"

/codon_start=1

/evidence=experimental

/product="BETA-1,2-N-ACETYLGLUCOSAMINYLTRANSFERASE I"

/protein_id="CAB70462.1"

/db_xref="GI:6779206"

/translation="MRGNKFCFLRLYLVLVAALAFIYIOMRLFATQSEKVRDRAAIE

AENHCTQSRLLDIKISQQGRVVALEQMKHQDQECRQLRALVQDLQSKGKGLIGD

VOMPVAAVVMAACSTDIYERTIKSLTKYTSVASKYPLFISQDGSNPDVKLALSYG

QJTYMQLDYEYVHTERPGELVAVYKIAHYKWLDDLPFKHNFYSRVIILEDMDIEIA

DEFDPEAGATLLDRDKSMAISSWNDNGQRFVQDPDALYSDSDFRGLGWLKSTAA

SELSPKWPKAYWDDWLKFNHRGRQFIRPEVCRIYNFGEHSSSLGQFQKYLEPIKL

NDVQVWKSMDLSYLLDENYVHFGLVKKAKPIHGADAVLKAFNIDGVRIOYRDQL

DPEDTARQFGIEFWKDGVPRAAYKGIWVFRFQTSRRVFLVSPDLSRLQGVEDT"

<80..>139

/gene="CGL"

CDS

/function="MEMBRANANKER (AMINO ACID 10-29)"
/standard_name="MEMBRANANKER EINES TYP II GOLGI-PROTEINS"
/note="DURCH VERGLEICH MIT TIERISCHEN GNTI-SEQUENZEN
IDENTIFIZIERT"
/codon_start=1
/product="HYDROPHOBER AMINO ACIDBEREICH IN GNTI"
/protein_id="CAB70463.1"
/db_xref="GI:6779207"
/translation="LRYLLVVAALAFIYIOMRLF"
1394..1655

3'UTR

ORIGIN

Query Match 72.3%; Score 1255; DB 6; Length 1669;

Best Local Similarity 89.1%; Pred. No. 2.3e-286;

Matches 1412; Conservative 0; Mismatches 140; Indels 33; Gaps 4;

Qy 126 GATGAGAGGGAACAAGTTTTCGTGTGATTTCCGGTACTCTCTCATCTTGGCTGCTGTCGC 185

Db 52 GATGAGAGGGAACAAGTTTTCGTGTGATTTACGGTACCTCTCTCGTGGCTGCTGTCGC 111

Qy 186 CTTTCATCTACACAGATGCGGCTTTTTCGACACAGTCAGAAATATGCAGATCGCTTGC 245

Db 112 CTTTCATCTACATACAGATGCGGCTTTTTCGACACAGTCAGAAATATGTAGACGCTTGC 171

Qy 246 TGCTGCAATTGAAGCAGAAATCATTTGTACAGCCAGACAGATGCTTTATTGACCAGAT 305

Db 172 TGCTGCAATTGAAGCAGAAATCATTTGTACAAAGTCAGACCAGATTTGCTTATTGCAAGAT 231

Qy 306 TAGCTGACAGCAAGGAGAGATAGTTCCTTGAAGACAAATGAAGCTCAGGACCAGGA 365

Db 232 TAGCAGCAGCAAGGAGAGATAGTTCCTTGAAGACAAATGAAGCTCAGGACCAGGA 291

Qy 366 GTGCCGCAATTAAGGGCTCTTGTTCAGGATCTTTGAAAGTAAGGCATATAAAGTTGAT 425

Db 292 GTGCCGCAATTAAGGGCTCTTGTTCAGGATCTTTGAAAGTAAGGCATATAAAGTTAAT 351

Qy 426 CGGAAATGTACAGATGCCAGTGGCTGTGTAGTTGTTATGGCTTGCAATGGGCTGATTA 485

Db 352 CGGAGATGTGCAGATGCCAGTGGCAGCTGTAGTTGTTATGGCTTGCAATGGCTGATTA 411

Qy 486 CTTGAAAGAGACTATTAAATCCATCTTAAATACCAATATCTGTGGTCAAAATATCC 545

Db 412 CTTGAGAGGAGCTATTAAATCCATCTTAAATACCAATATCTGTGGTCAAAATATCC 471

Qy 546 TCTTTTCATATCCAGGATGGATCACATCTCATGTTCAGGAGCTTGTCTTGAGCTATGA 605

Db 472 TCTTTTCATATCCAGGATGGATCAATCTCATGTTCAGGAGCTTGTCTTGAGCTATGG 531

Qy 606 TCAGCTGACGTATATGCAGCAGCTTGGATTTTGAACCTGTGCATATCTGAAAGCCAGGGA 665

Db 532 TCAGCTGACGTATATGCAGCAGCTTGGATTTTGAACCTGTGCATATCTGAAAGCCAGGGA 591

Qy 666 GCTGATTGCATATCAAAATTCGACGTCATTACAGTGGGCGATGGATCAGCTGTTTGA 725

Db 592 ACTGGTTGCATATCAAAATTCGACGTCATTACAGTGGGCGATGGATCAGCTGTTTGA 651

Qy 726 CAAGCATAAATTTAGCCGTGTTATCATACTAGAAGATGATGGAATTTGCCCTGATTT 785

Db 652 CAAGCATAAATTTAGCCGTGTTATCATACTAGAAGATGATGGAATTTGCTGCTGATTT 711

Qy 786 TTTTGACTTTTTTGAGCTGGAGCTACTCTTCTTGACAGACAAGTCGATTTATGGCTAT 845

Db 712 TTTTGACTTTTTTGAGCTGGAGCTACTCTTCTTGACAGACAAGTCGATTTATGGCTAT 771

Qy 846 TTCTCTTGGAAATGACAAATTCGACGTTGTTCCAGATCTTATGCTCTTTTACCG 905

Db 772 TTCTCTTGGAAATGACAAATTCGACGTTGTTCCAGATCTTATGCTCTTTTACCG 831

Qy 906 CTCAGATTTTTTCCCGTCTTGGATGATGCTTTTCAAAATCTACTTGGGAGCAATATTC 965

Db 832 CTCAGATTTTTTCCCGTCTTGGATGATGCTTTTCAAAATCTACTTGGTCCGAACTATC 891

Qy 966 TCCAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTTAAGACTCAAGAGAAATCACAGAGG 1025

Db	232	TAGCCAGCAGCAAGGAGAGTAGTGCTCTTTGAAGAACAAATGAAGCATCAGGACCGAGA	291
Qy	366	GTGCCGACAATTAAGGGCTCTTGTTTCAGGATCTTTGAAAGTAGAAGGCATAAAAAGTTGAT	425
Db	292	GTGCCGCCAATTAAGGGCTCTTGTTTCAGGATCTTTGAAAGTAGAAGGCATAAAAAGTTAAT	351
Qy	426	CGGAAATGTTACAGATGCCAGTGGCTGCTGTAGTTGTTATGCGCTTGCATTCGGGCTGATTA	485
Db	352	CGGAGATGTCAGATGCCAGTGGCAGCTGTAGTTGTTATGCGCTTGCAGTCGTACTGACATTA	411
Qy	486	CCTGGAAAAAGCATTTAAATCCATCTTAAATAACCAAAATATCTGTTGCGGTCAAAAATATCC	545
Db	412	CCTGGAGAGGACTATTAAATCCATCTTAAATAACCAAAATATCTGTTGTCATCAAAATATCC	471
Qy	546	TCCTTTTCATATCCCAGGATGGATCACATCCTGATGTGAGGAAGCTTGCTTTGAGCTATGA	605
Db	472	TCCTTTTCATATCCCAGGATGGATCAAATCCTGATGTAAAGAAAGCTTGCTTTGAGCTATGG	531
Qy	606	TCAGCTGACGTAATGAGCAGACTTTGGATTTTGAACCTGTGCATACCTGAAAGACCGGGGA	665
Db	532	TCAGCTGACGCTATATGAGCAGCACTTGGATTTATGAACCTGTGCATACCTGAAAGACCGGGGA	591
Qy	666	GCTGATTTGCATACTACAAAATTCGACGCTCAATTACAAGTGGGCACTGGATCAGCTGTTTTA	725
Db	592	ACTGGTTGCATACTACAAGATTTGCAGCTCAATTACAGTGGGCACTGGATCAGCTGTTCAT	651
Qy	726	CAAGCATAATTTTAGCCGCTGTTATCATATAGAAAGATGATGGAATTCGCCCTGATTT	785
Db	652	CAAGCATAATTTTAGCCGCTGTTATCATATAGAAAGATGATGGAATTCGCTGATTT	711
Qy	786	TTTTGACTTTTTTGAAGCTGAGGCTACTCTTCTTTGACAGACAAAGTCGATATAGGCTAT	845
Db	712	TTTTTGACTATTTTGAGGCTGAGGCTACTCTTCTTTGACAGACAAAGTCGATATAGGCTAT	771
Qy	846	TTCTTCTTGGAAATGACAAATGCAAAATGCAAGTTTGTCCAAAGATCCTTATGCTCTTTACCG	905
Db	772	TTCTTCTTGGAAATGACAAATGCAAAAGGCAAGTTTGTCCAAAGATCCTGATGCTCTTTACCG	831
Qy	906	CTCAGATTTTTTTCCCGCTCTTGATGGATGCTTTCAAAATCTACTTTGGGACGAAATATATC	965
Db	832	CTCAGATTTTTTTCTGCTCTTGGATGGATGCTTTCAAAATCAACTTTGGTCCGAACTATC	891
Qy	966	TCCAAAGTGGCCAAAGGCTTACTCGGACGACTGGCTTAAGACTCAAGAGAAATCACAGAGG	1025
Db	892	TCCAAAGTGGCCAAAGGCTTACTCGGATGACTGGCTTAAGGCTGAAGAAATCACAGAGG	951
Qy	1026	TCGACAAATTTATTCGCCCAAGATTTGACAGAAATATATATTTGGTGAGCATGGTCTAG	1085
Db	952	TCGACAAATTTATTCGCCCAGAAAGTTTGACAGAACTGACAAATTTTGGTGAGCATGGTCTAG	1011
Qy	1086	TTTGGGGCAGTTTTTCAAGCAGATATCTTGAGCCAAATTAACCTAAATGATGTCCAGGTTGA	1145
Db	1012	TTTGGGGCAGTTTTTAAAGCAGATATCTTGAGCCAAATTAAGCTAAATGATGTCCAGGTTGA	1071
Qy	1146	TTGGAAGTCAATTGAACCTTAGTTACTCTTTTGGAGGACAATTAACGTAAGCAACTTTGTGTA	1205
Db	1072	TTGGAAGTCAATTGAACCTTAGTTACTCTTTTGGAGGACAACACTATGTGTAAGCAACTTTGGCGA	1131
Qy	1206	CTTGTTAAAAAGGCTAAGCCCATCCATGGAGCTGATGCTGTCTTTGAAAGCAATTAACAT	1265
Db	1132	CTTGTTAAAAAGGCTAAGCCCATCCACGGAGCTGATGCTGTCTTTTGAAGCAATTAACAT	1191
Qy	1266	AGATGGTGAATGTCGCTATTTACGATACAGAGATCACTAGACTTTTGAAGATATCGCACCGCA	1325
Db	1192	AGATGGTGAATGTCGCTATTTACGATACAGAGACCAACTAGACTTTTGAAGATATCGCTCGACA	1251
Qy	1326	ATTTTGGCATTTTTTGAAGAAATGGAAGGATGGTGTACCAAGCTGCAGCATATAAAGGAATAGT	1385
Db	1252	GTTTTGGCATTTTTTGAAGAAATGGAAGGATGGTGTACCAAGGAGCATATAAAGGATAGT	1311
Qy	1386	AGTTTTCCGGTACCAAAAGCTCCAGACGTGTATTCCTTTTGGCCCAATGATTCGTTTCAACA	1445
Db	1312	AGTTTTCCGGTTCAAAATCTAGACGTTGTTCTTCTTTTCCCTGATTCCTTTTCAACA	1371

Qy	1446	ACTCGGAATTGAAGATACCTTAAACAAGATATGATGTGCAGAGCCCGGCAAAATTTTGA	1505
Db	1372	ACTTGGAGTTCAAGATACCTTACGCAAGATATGATT---GGAGCCTGAGCAACAATTTAGA	1428
Qy	1506	CTTATTGGGTAGGATGCATCGAGCTGACACTAAACCATGATTTT-----ACCAGTTAC	1558
Db	1429	CTTATTTTGGTAGGATACATTTTGAAGAGCTGACACGAAGAATGATGACTACACAGTAGCTAC	1488
Qy	1559	ATACAACGTTTTAAATGTTTATACGGAGAGCTCACGTCTTAGTGTGTAAGGATATCGGC	1618
Db	1489	ATGCAACATTTTAAATGTTTAAATGGAAGGAACCCACTGCT-----T	1527
Qy	1619	TTCTTAGTATTGGATGAATCATCAACAACCTATTATTTTAAAGTTTCAGAACATAAAG	1678
Db	1528	ATTGTTGGAATGGATGAATCATCACCAATCCTATTA--TTCAAGTTTACAACAATAAAG	1588
Qy	1679	AGGAATGCTAGCCCTGTAAGACTTA	1703
Db	1586	AGGAATGTTGCCCTATATAAACAA	1610
RESULT 11			
AR432618			
LOCUS			
AR432618			
DEFINITION			
Sequence 1 from patent US 6653459.			
ACCESSION			
AR432618.1			
VERSION			
AR432618.1			
KEYWORDS			
Unknown.			
SOURCE			
Unknown.			
ORGANISM			
Unclassified.			
REFERENCE			
1 (bases 1 to 1669)			
AUTHORS			
Von Schaeuwen,A.			
TITLE			
Plant GntI sequences and the use thereof for the production of			
plants having reduced or lacking N-acetyl glucosaminyl transferase			
I(GntI) activity			
JOURNAL			
Patent: US 6653459-A 1 25-NOV-2003;			
FEATURES			
source			
1..1669			
Location/Qualifiers			
/organism="unknown"			
/mol_type="genomic DNA"			
ORIGIN			
Query Match			
Best Local Similarity 72.3%; Score 1255; DB 6; Length 1669;			
Matches 1412; Conservative 89.1%; Pred. No. 2.3e-296;			
Indels 33; Gaps 4			
Qy	126	GATGAGAGGAACAAGTTTTGCTGTGATTTCCGGTACCTCTCATCTTGGCTGCTGTGCG	185
Db	52	GATGAGAGGAACAAGTTTTGCTTTGATTTACGGTACCTTCTCGTGTGGCTGCTCTCGC	111
Qy	186	CTTCATCTACACAGATGCGGCTTTTTCGACACAGTCAGNATATGCAGATCGCCTGCG	245
Db	112	CTTCATCTACATACAGATGCGGCTTTTTCGACACAGTCAGNATATGTAGACCGCCTTGC	171
Qy	246	TCCTCAATTGAACAGAGAAATCATTTGTACAAGCCAGACCAATGCTTATTGACCAGAT	305
Db	172	TCCTCAATTGAACAGAGAAATCATTTGTACAAGTCAGACCAATGCTTATTGACAAGAT	231
Qy	306	TAGCCTGCAGCAAGGAAGAAATAGTTGCTCTTGAAGAACAAATGAAGCGTCAGACCAAGA	365
Db	232	TAGCCAGCAGCAAGGAAGAGTAGTAGTCTTGAAGAACAAATGAAGCATCAGGACCAAGA	291
Qy	366	GTGCGGCAATTAAGGGCTCTTGTTCAGGATCTTGAAGTAGGCGATATAAAGATTGAT	425
Db	292	GTGCGGCAATTAAGGGCTCTTGTTCAGGATCTTGAAGTAGGCGATATAAAGATTGAT	351
Qy	426	CGGAATGTACAGATGCCAGTGGCTGCTGTAGTGTGTTATGGCTTGCAATCGGGCTGATTA	485
Db	352	CGGAATGTACAGATGCCAGTGGCGTGTGTTATGGCTTGCAATCGTACTGACTA	411
Qy	486	CCTGGAAAAGACTATTAAATTCATCTTAAATACCAAAATATCTGTTGGGTGAAAATATCC	545

Db 412 CCTGGAGGACTATTAATCCATCTTAAATACCAACATCTGTTGCATCAAAATATCC 471
Qy 546 TCTTTTCATATCCAGGATGATCACATCCCTGATGTGAGAAAGTCTGTTTGAGCTATGA 605
Db 472 TCTTTTCATATCCAGGATGATCAAAATCCCTGATGTGAAGAAAGCTGCTTTGAGCTATGG 531
Qy 606 TCAGCTGAGTATATGAGCACTTTGGATTTTGAACCTGTGCATATCAGAAAGCAGGGGA 665
Db 532 TCAGCTGAGTATATGAGCACTTTGGATTTTGAACCTGTGCATATCAGAAAGCAGGGGA 591
Qy 666 GCTGATTGCATACAAAAATTCACGCTCATTAACAAGTGGCATTTGGATCAGCTGTTTTTA 725
Db 592 ACTGTTGCTACTACAAGATTCACGCTCATTAACAAGTGGCATTTGGATCAGCTGTTTCA 651
Qy 726 CAAGCATAAATTTAGCCGTGTTATCATACTAGAAAGATATGAAATATGCCCCCTGATTT 785
Db 652 CAAGCATAAATTTAGCCGTGTTATCATACTAGAAAGATATGAAATATGCTGTGATTT 711
Qy 786 TTTTGTCTTTTTCAGGCTGGCTACTCTTCTTGACAGACAAAGTCGATTATGGCTAT 845
Db 712 TTTTGTCTTTTTCAGGCTGGCTACTCTTCTTGACAGACAAAGTCGATTATGGCTAT 771
Qy 846 TTTTGTCTTTTTCAGGCTGGCTACTCTTCTTGACAGACAAAGTCGATTATGGCTAT 905
Db 772 TTTTGTCTTTTTCAGGCTGGCTACTCTTCTTGACAGACAAAGTCGATTATGGCTAT 831
Qy 906 CTCAGATTTTTCCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 965
Db 832 CTCAGATTTTTCCTGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 891
Qy 966 TCCAAAGTGGCCAAAGCTTACTGGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1025
Db 892 TCCAAAGTGGCCAAAGCTTACTGGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 951
Qy 1026 TCGACAAATTTATTCGCCAGAAAGTTTGCAGAAATATTAATTTTGGTGAGCATGTTCTAG 1085
Db 952 TCGACAAATTTATTCGCCAGAAAGTTTGCAGAAATATTAATTTTGGTGAGCATGTTCTAG 1011
Qy 1086 TTTGGGCGAGTTTTCAGAGCATATCTTGGAGCAATTTACGTGAAACACTTTGCTGTA 1145
Db 1012 TTTGGGCGAGTTTTCAGAGCATATCTTGGAGCAATTTACGTGAAACACTTTGCTGTA 1071
Qy 1146 TTTGAGTCAATGACCTTACTTACTTTTGGAGCAATTTACGTGAAACACTTTGCTGTA 1205
Db 1072 TTTGAGTCAATGACCTTACTTACTTTTGGAGCAATTTACGTGAAACACTTTGCTGTA 1131
Qy 1206 CTTGGTTAAAAAGGCTAAGCCCATCCATGGAGCTGATGCTCTTTGAAAGCATTTAAT 1265
Db 1132 CTTGGTTAAAAAGGCTAAGCCCATCCATGGAGCTGATGCTCTTTGAAAGCATTTAAT 1191
Qy 1266 AGATGGTGAATGCGTATTCAGTACAGAGATCAACTAGACTTTGAAATATCGCACGGCA 1325
Db 1192 AGATGGTGAATGCGTATTCAGTACAGAGATCAACTAGACTTTGAAATATCGCTCGACA 1251
Qy 1326 ATTTGGCATTTTGAAGATGAGGATGTTGTTACCGCTGACATATAAGGAATAGT 1385
Db 1252 GTTTGGCATTTTGAAGATGAGGATGTTGTTACCGCTGACATATAAGGAATAGT 1311
Qy 1386 AGTTTTCGGTACCAACGCTCCAGACGTTATTCCTTGTGGCCATGATTCGCTTCAACA 1445
Db 1312 AGTTTTCGGTACCAACGCTCCAGACGTTATTCCTTGTGGCCATGATTCGCTTCAACA 1371
Qy 1446 ACTCGGAATGAAGATATCTTAACAAAGATATGATTTGAGGAGCCCGGGCAAAATTTTGA 1505
Db 1372 ACTTGGAGTTGAAGATATCTTAACAAAGATATGATTTTGAAGGCTGAGCAACATTTAGA 1428
Qy 1506 CTTATTTGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1558
Db 1429 CTTATTTGGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1488
Qy 1559 ATACAAGCTTTTAATGTTATACGGAGGAGCTCACTGTTCTAGTGTGAAAGGATATCGGC 1618

Db 1489 ATGCAACATTTTAATGTTAAATGGAAGGAACCCACTGCT-----T 1527
Qy 1619 TTCTTTAGTATTGGATGAATCATCAACACACAACTTATTTTAACTGTTTCAAGACATAAAG 1678
Db 1528 ATTGTTGGAATGGATGAATCATCAACACATCCTCTATTTA--TTCAAGTTTACAAACATAAAG 1585
Qy 1679 AGAAAAATGTAGCCCTCTAAAGACTA 1703
Db 1586 AGAAAAATGTAGCCCTCTAAAGACTA 1610

BT013875 1719 bp mRNA linear PLN 11-MAY-2004
Lycopersicon esculentum clone 132855F, mRNA sequence.

BT013875
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BT013875.1 GI:47105290
FLI CDNA.
Lycopersicon esculentum (Solanum lycopersicum)
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

1 (bases 1 to 1719)
Kirkness, E.F., Wang, W. and Vazeille, A.
Direct Submission
Submitted (11-MAY-2004) The Institute for Genomic Research, 9712
Medical Center Drive, Rockville, MD 20850, USA

FEATURES
Location/Qualifiers
1..1719
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="132855F"
/tissue_type="fruit"
/note="TZXH109"

ORIGIN

Query Match 71.3%; Score 1239.2; DB 15; Length 1719;
Best Local Similarity 90.8%; Pred. No. 1.7e-292;
Matches 1356; Conservative 0; Mismatches 128; Indels 10; Gaps 3;

Qy 126 GATGAGAGGGAACAAGTTTGTCTGTGATTTCCGGTACCTCTCATCTTGGCTGCTGTCGC 185
Db 86 GATGAGAGGGAACAAGTTTGTCTGTGATTTACGGTACCTCTCTCTGCTGCTGCTCGC 145
Qy 186 CTTTCATCTACACAGATCGGCTTTTTCGCACACAGTCAAGATATGACATCGCCTTGC 245
Db 146 CTTTCATCTACATACAGATCGGCTTTTTCGCACACAGTCAAGATATGACATCGCCTTGC 205
Qy 246 TGTGCAATTTGAAGCAAGAAATCATTTGTACAAAGCCAGACAGATTTGCTTATTGACCATGAT 305
Db 206 TGTGCGAATTTGAAGCAAGAAATCATTTGTACAAAGTCAAGTCTAGTTGCTTATTGACCAAGAT 265
Qy 306 TACCTGTCAGCAGGGAAGATAGTCTCTTTGAAGAACAAATCAAGCTCAGACACAGGA 365
Db 266 TACCCAGCAGCAAGGAAGTAGTGTCTTTGAAGAACAAATCAAGCTCAGACACAGGA 325
Qy 366 GTGCCGACAAATTAAGGGCTCTTGTTCAGGATCTTTGAAGTAAGGGCATATAAAGTTGAT 425
Db 326 GTGCCGACAAATTAAGGGCTCTTGTTCAGGATCTTTGAAGTAAGGGCATATAAAGTTGAT 385
Qy 426 CGGAATTTACAGATCCAGTGGCTGCTGTAGTTGTTATGGCTTGCATCGGCTGATTA 485
Db 386 CGGAATTTGTCAGATCCAGTGGCTGCTGTAGTTGTTATGGCTTGCATCGGCTGATTA 445
Qy 486 CTTGGAAAGACATTTAAATCCATCTTAAATACCAAAATATCTGTTGCGTCAAAATATCC 545
Db 446 CTTGGAGAGACATTTAAATCCATCTTAAATACCAAAATATCTGTTGCGTCAAAATATCC 505
Qy 546 TCTTTTTCATATCCAGGATGGATCAGATCCTGATGTGAGAAAGCTTGTCTTTGAGCTATGA 605
Db 506 TCTTTTTCATATCCAGGATGGATCAGATCCTGATGTGAGAAAGCTTGTCTTTGAGCTATGA 565

Qy	606	TCAGCTGACGTATATGCAGCACTTGGATTTTGGAACTGTGCATACCTGAAAGACACGAGGA	665
Db	566	TCAACTGACGTATATGCAGCACTTGGATTTTGGAACTGTGCATACCTGAAAGACACGAGGA	625
Qy	666	GCTGATGTGCATCTAATAAAATTGCAACGTCAATACAAAGTGGGCATTTGGATCAGCTGTTTAA	725
Db	626	ACTGGTTGCATCTAATAAAATTGCAACGTCAATACAAAGTGGGCATTTGGATCAGCTGTTTCA	685
Qy	726	CAAGCAATAATTTAGCCGTGTATCATCTAGAAAGATGATATGGAATTTGCCCTTGATTT	785
Db	686	CAAGCAATAATTTAGCCGTGTATCATCTAGAAAGATGATATGGAATTTGCCCTTGATTT	745
Qy	786	TTTTTGACTTTTTTGAGGCTGGAGCTACTCTCTTTTGACAGAGACAAGTCGATTATGGCTAT	845
Db	746	TTTTTGACTACTTTTGAGGCTGGAGCTACTCTCTTTTGACAGAGACAAGTCGATTATGGCTAT	805
Qy	846	TTCTTCTTGGAAATGAATAATGGAATAATGCAAGTTTGTCCAAAGATCCTTATGTCTCTTTACCG	905
Db	806	TTCATCTTGGAAATGAATAATGGAATAATGCAAGTTTGTCCAAAGATCCTTATGTCTCTTTACCG	865
Qy	906	CTCAGATTTTTTTCCGGTCTTGGATGGATGCTTTTCAAAATCTACCTTTGGGAGCAATATC	965
Db	866	CTCAGATTTTTTTCCGGTCTTGGATGGATGCTTTTCAAAATCTACCTTTGGGAGCAATATC	925
Qy	966	TCCAAAGTGCCAAAGGCTTACTTGGGACGACTGGCTAAAGACTCAAGAGAGAATCACAGAGG	1025
Db	926	TCCAAAGTGCCGGAAGGCTTACTGGGATGACTGGCTGAGGCTCAAGAGAAAATCACAGAGG	985
Qy	1026	TCGACAAATTTATTCGCCGAGAAAGTTTGCAGAAATATAATAATTTTGGTGGAGCATGGTCTTAG	1085
Db	986	TCGACAAATTTATTCGCCGAGAAAGTTTGCAGAAAGTATAATAATTTTGGTGGAGCATGGTCTTAG	1045
Qy	1086	TTTGGGCGAGTTTTTCAAGCAGTATCTTGAGCCAAATTAACATAAATGATGCTCCAGAGTCA	1145
Db	1046	TTTGGGCGAGTTTTTAAAGCAGTATCTTGAGCCAAATTAACATAAATGATGCTCCAGAGTCA	1105
Qy	1146	TTGGAAGTCAATGGACCTTAGTTACCTTTTGGAGGACAAATACGTGAAAAACATTTTGGTGA	1205
Db	1106	TTGGAAGTCAATGGACCTTAGTTACCTTTTGGAGGACAAATACGTGAAAAACATTTTGGGA	1165
Qy	1206	CTTGGTTAAAAAGGCTAAGCCCATCCATGGAGCTGATGCTGTCTTGAAGAGCATTTAAACAT	1265
Db	1166	CTTGGTTAAAAAGGCTAAGCCCATCCATGGAGCTGATGCTGTCTTGAAGAGCATTTAAACAT	1225
Qy	1266	AGATGGTGATGTGGTATTCAGTACAGAGATCAACTAGACTTTTGAAATAATCCACGGCA	1325
Db	1226	AGATGGTGATGTGGTATTCAGTACAGAGATCAACTAGACTTTTGAAATAATCCACGGCA	1285
Qy	1326	ATTTGGCATTTTTCAGAAATGGAAGGATGGTGACCACTGGACGATATAAAGGAATAGT	1385
Db	1286	GTTTGGAAATTTTGAAGATGGAAGGATGGTGACCACTGGACGATATAAAGGAATAGT	1345
Qy	1386	AGTTTTCCGGTACCAACCGTCCAGACGTGATATTCCTGTGTGGCCATGATTCGCTTCAACA	1445
Db	1346	AGTTTTCCGGTACCAACCGTCCAGACGTGATATTCCTGTGTGGCCATGATTCGCTTCAACA	1405
Qy	1446	ACTCGGAATTGAAGATCTTAAAGAGATATGATGTGCAGGAGCCCGGGCAAAATTTTGA	1505
Db	1406	ACTTGGAGTTGAAGATCTTAAAGAGATATGATGTGTGCAGGAGCCCGGGCAAAATTTTGA	1465
Qy	1506	CTTATTTGGTGGATGCAAT-----CGAGCTGACACTAAACCATGATTT--TACCAAGTTAC	1568
Db	1463	CTTATTTGGCAGGATACATTTTGAAGAGCTGGCACGAAGAGCTTGATTAGCAGTAGTGTGC	1522
Qy	1559	ATACAAAGCTTTTAATGTTTATACGAGGAGCTCAGCTGTGTTCTAGTGTGAAAGGAT	1612
Db	1523	ATACAAACATTTTAATGTTTAAATGGAAGGAACTCACTGCTACTGTGTGGAGTGGAT	1576

Db 184 GTCAGGACCAGGAGTCGCGGCAATTAAGGGCTCTTGTTCCAGGATCTTTGAAAGTAAGGGCA 243
Qy 413 TAAAAAGTTGATCGGAAATGTACAGATGCGAGTGGCTGCTGTGTGTTGTTATGCTTTGCA 472
Db 244 TAAAAAGTTAATCGGAGATGTGCAGATGCGAGTGGCGGCTGTAGTTGTTATGCTTTGCA 303
Qy 473 ATCGGGCTGATTACTCGGAAAGACTATTAATCCATCTTTAAATACCAATATCTGTTG 532
Db 304 GTGCTGTGACTACTCGGAGAAGACTATTAATCCATCTTTAAATACCAATATCTGTTG 363
Qy 533 CGTCAAAATATCTCTTTTTCATATCCAGGATGGATCAATCTCTGTATGTCAGGAAGCTTG 592
Db 364 CATCAAAATATCTCTTTTTCATATCCAGGATGGATCAATCTCTGTATGTCAGGAAGCTTG 423
Qy 593 CTTTGACTATGATCAGCTGACGTATATGAGCACTTGGATTTTGAACCTGTGCACTAGT 652
Db 424 CTTTGACTATGATCAGCTGACGTATATGAGCACTTGGATTTTGAACCTGTGCACTAGT 483
Qy 653 AAAGACCAGGGAGCTGATTGCTACTACTACAAATTTGCAAGTTCATTAAGTGGGCAATTGG 712
Db 484 AAAGACCAGGGAGAACTGGTTGCATACTACAAAGTTGCAAGTTCATTAAGTGGGCAATTGG 543
Qy 713 ATCAGCTGTTTTCACAGCAATAATTTAGCCGCTGTTATCATACTAGAGATCATATGGAAA 772
Db 544 ATCAGCTGTTTTCACAGCAATAATTTAGCCGCTGTTATCATACTAGAGATCATATGGAAA 603
Qy 773 TTGCCCTGATTTTGTGACTTTTGGAGCTGAGCTACTCTTCTTGAAGAGACAGCAAGT 832
Db 604 TTGCTGCTGATTTTGTGACTTTTGGAGCTGAGCTACTCTTCTTGAAGAGATAAGT 663
Qy 833 CGATTATGGCTATTCTCTTGGAAATCAATGACAAATGCAAGTTCAGTTCCTCAAGATCCTT 892
Db 664 CGATTATGGCTATTCTCTTGGAAATCAATGACAAAGGAGTTCGTCAGAGATCCCT 723
Qy 893 ATGCTCTTTACCGCTCAGATTTTTCCTCCGCTCTTGGATGGATGCTTTTCAAAATCTACTT 952
Db 724 ATGCTCTTTACCGCTCAGACTTTTTCCTGCTTGGATGGATGCTTTTCAAAATCAACTT 783
Qy 953 GGGACGAATTTATCTCAAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTAAGACTCAAG 1012
Db 784 GGTCCGAATCTATCTCAAAAGTGGCCAAAGGCTTACTGGGATGACTGGCTAAGGCTCAAG 843
Qy 1013 AGAATCACAGAGTTCGCAATTTATTCGCCAGAGTTTCAGAACATATAATTTTGGTG 1072
Db 844 AAAATCACAGAGTTCGCAATTTATTCGCCAGAGTTTTCAGAACATATAATTTTGGTG 903
Qy 1073 AGCATGTTCTAGTTTGGGGCAGTTTTCAGCAGTATCTTGAAGCCAAATTAACCTAAATG 1132
Db 904 AGCATGTTCTAGTTTGGGGCAGTTTTCAGCAGTATCTTGAAGCCAAATTAAGCTAAATG 963
Qy 1133 ATGTCAGGTTGATTGGAAGTCAATGGAACCTTAGTTACCTTTTGGAGGACAATTAAGCTGA 1192
Db 964 ATGTCAGGTTGATTGGAAGTCAATGGAACCTTAGTTACCTTTTGGAGGACAATTAAGCTGA 1023
Qy 1193 AACACTTTGTGTGACTGTTTAAAGAGCTAAGCCATCCATGAGCTGATGCTGCTTGA 1252
Db 1024 AACACTTTGTGTGACTGTTTAAAGAGCTAAGCCATCCATGAGCTGATGCTGCTTGA 1083
Qy 1253 AAGCATTTAATAGATGTTGCTGATGCTGCTATTCAGTACAGAGATCAACTAGACTTTGAAA 1312
Db 1084 AAGCATTTAATAGATGTTGCTGATGCTGCTATTCAGTACAGAGATCAACTAGACTTTGAAA 1143
Qy 1313 ATATCGCACGGCAATTTGGCAATTTTGAAGAAAGGATGGTGATACCAAGTGCAGCAT 1372
Db 1144 ATATCGCTCGGCAGTTTGGCAATTTTGAAGAAAGGATGGTGATACCAAGTGCAGCAT 1203
Qy 1373 ATAAAGGAATAGTAGTTTCCGGTACCAGCTCCAGCTGATCTCTGTTGGCCATG 1432
Db 1204 ATAAAGGAATAGTAGTTTCCGGTACCAGCTCCAGCTGATCTCTGTTGGCCATG 1263
Qy 1433 ATTGCTCTTCAACACTCGGAATTTGAAGATATCTTAACAAAGATATGATTCGAGGAGCCGG 1492
Db 1264 ATTCTCTTCACAACACTTGGAGTTGAAGATACGTAGCGAAGATATGATTTTGGTGCCTGA 1321

Qy 1493 GCAAAATTTTGCATTTATTTGGGTAGCATGCAT-----CGAGCTGACACTAAA--CCATGA 1545
Db 1322 GCAACACTTTTAGACTTTATTTGGTAGGATACATTTGAAGGAGCTGCACACGAAAAGCATGAC 1381
Qy 1546 TTTTACCAGTTACATCAACAGTTTAAATGTTATACCGAGGAGCTCACTGTTCTAGTGTG 1605
Db 1382 TACCAGTAGCTACATGCAACATTTTAAATGTTATGAAGGAACCTCACTGCCTACTGTTG 1441
Qy 1606 AAGGGAT 1612
Db 1442 AATGGAT 1448

RESULT 14
STU249880

LOCUS

DEFINITION STU249880 1499 bp mRNA linear PLN 15-APR-2005
Solanum tuberosum mRNA for N-acetylglucosaminyltransferase I (GntI
gene), clone A8.

ACCESSION

AJ249880

VERSION

AU249880.1 GI:18076143

KEYWORDS

GntI gene; N-acetylglucosaminyltransferase I.

SOURCE

Solanum tuberosum (potato)

ORGANISM

Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum.
1
Wenderoth, I., Tjaden, J. and von Schaewen, A.
Isolation and characterization of different plant
N-acetylglucosaminyltransferase I (GntI) cDNA sequences: Functional
analyses in the Arabidopsis cgl mutant, and in potato and tobacco
antisense plants
Unpublished

JOURNAL

REFERENCE 2 (bases 1 to 1499)

AUTHORS

von Schaewen, A.

TITLE

Direct Submission

JOURNAL

Submitted (28-SEP-1999) von Schaewen A., Pflanzenphysiologie,
Universitaet Osnabrueck, FB5, Barbarastr. 11, D-49076 Osnabrueck,
GERMANYFEATURES
source

1. .1499
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Desiree"
/db_xref="taxon:4113"
/clone="A8"
/tissue_type="tuber"
/note="Sink"
1. .1499
/gene="GntI"
37. .1158
/gene="GntI"
/EC_number="2.4.1.101"
/function="initiates complex-type glycans in the Golgi"
/codon_start=1
/product="N-acetylglucosaminyltransferase I"
/protein_id="CAC80699.1"
/db_xref="GI:18076144"
/db_xref="GOA:O8VX54"
/db_xref="InterPro:IPR004139"
/db_xref="UniProt/TREMBL:O8VX54"
/translation="MKRQDECRQRLVQDLLESKIKKLGIVQMPVAIVVMACSR
ADYLERIKSILKYQTSVASKYPLFTISQGSNPDVRKLALSYDLTYMQHLDPEVHT
EKPGLVAYKIARHYKWDQLDFHKHNSRVILLEDDEIAADFFDYFEAGATLLDR
DKS IMAISSNDNGQRFQVDALVRSDFPGLGWMLSKSTWSELSKPWKAYWDDW
LRKKNHRGQFTHPEVCTYNGEHSLSIGQFKQYLEPIKLNDVQDWKMDLSYL
LEDNYVKKFDLVKKAKPIHGADVLKAFNIGDVRIGRIQVRDQLDFAIARQFGIFBEW
KDGVPRAAYKGIWVFRFQTSRRRVFLVAPDSLRLQLGVEDT"

ORIGIN

Query Match 62.1%; Score 1078; DB 15; Length 1499;
Best Local Similarity 87.5%; Pred. No. 5.3e-253;

Matches 1237; Conservative 0; Mismatches 145; Indels 32; Gaps 4;	
Qy	330 TGCTCTTGAAGAACAAATGAAGCTCAGGACCAGAGTGGCGACAATTAAGGGCTCTTGT
Db	21 TGTTCCTCGGTAAACAAATGAAGCTCAGGACCAGAGTGGCGACAATTAAGGGCTCTTGT
Qy	390 TCAGGATCTTGAAGTAAAGGACATAAAAAAGTTGATCGGAAATGTACAGATGCCAGTGGC
Db	81 TCAGGATCTTGAAGTAAAGGACATAAAAAAGTTAATCGGAGATGTGCGAGATGCCAGTGGC
Qy	450 TGCTGTAGTTTATGGCTGCAATTCGGGCTGTATTAATCGAATTAATTAATCGAT
Db	141 AGCTGTAGTTTATGGCTGCAATTCGGGCTGTATTAATCGAATTAATTAATCGAT
Qy	510 CTTAAATACCAATATCTGTTCGCTCAAAATATCTCTTTTCATATCCAGGATGGATC
Db	201 CTTAAATACCAATATCTGTTCGCTCAAAATATCTCTTTTCATATCCAGGATGGATC
Qy	570 ACATCCTGATCTCAGGAAGCTTGTCTTGAGCTATGATCAGCTGAGCTATATGCGACACTT
Db	261 AATCCTGATGTAGAAGCTTGTCTTGAGCTATGATCAGCTGAGCTATATGCGACACTT
Qy	630 GGATTTTGAACCTGTGCATCTGAAAGACAGGAGGAGCTGATGATCACTACAAATTCG
Db	321 GGATTTTGAACCTGTGCATCTGAAAGACAGGAGGAGCTGATGATCACTACAAATTCG
Qy	690 ACCTCATACCAAGTGGGATTTGGATCAGCTGTTTACCAAGCATATTTAGCCGTGTAT
Db	381 ACCTCATACCAAGTGGGATTTGGATCAGCTGTTTACCAAGCATATTTAGCCGTGTAT
Qy	750 CATACTAGAAGATGATATGGAATTTGCCCTCGATTTTTCGACTTTTTCGAGGCTGGAGC
Db	441 CATACTAGAAGATGATATGGAATTTGCCCTCGATTTTTCGACTTTTTCGAGGCTGGAGC
Qy	810 TACTCTTCTGCACAGACAAGTCAATATGCTATTTCTTCTGGAATGACAATGGACA
Db	501 TACTCTTCTGCACAGACAAGTCAATATGCTATTTCTTCTGGAATGACAATGGACA
Qy	870 AATGAGTTTGTCCAGATCTTATGCTCTTTACCGCTCAGATTTTTCGCGTCTGG
Db	561 AAGGAGTTTGTCCAGATCTTATGCTCTTTACCGCTCAGATTTTTCGCGTCTGG
Qy	930 ATGAGTCTTTCACAAATCTACTTGGGAGCAATATCTCAAGTGGCCAAAGGCTTACTG
Db	621 ATGAGTCTTTCACAAATCTACTTGGGAGCAATATCTCAAGTGGCCAAAGGCTTACTG
Qy	990 GGACGACTGGCTAAGACTCAAGAGAAATCAGAGGTCGACAATTTATTCGCCAGAAAT
Db	681 GGATGACTGGCTAAGCTGAAAGAAATCAGAGGTCGACAATTTATTCGCCAGAAAT
Qy	1050 TTGCAGAACATATAATTTTGTGAGCATGTTCTAGTTTGGGCGAGTTTTCAGCAGTA
Db	741 TTGCAGAACATATAATTTTGTGAGCATGTTCTAGTTTGGGCGAGTTTTCAGCAGTA
Qy	1110 TCTTCAGCAATTAACATTAATGATGCTCCAGTTCATTTGGAAGTCAATGACCTTAGTTA
Db	801 TCTTCAGCAATTAACATTAATGATGCTCCAGTTCATTTGGAAGTCAATGACCTTAGTTA
Qy	1170 CTTTTCGAGGACAAATTAAGTGAACACATTTTGGTACCTTGGTTAAAAAGGCTTAAGCCAT
Db	861 CTTTTCGAGGACAAATTAAGTGAACACATTTTGGTACCTTGGTTAAAAAGGCTTAAGCCAT
Qy	1230 CCAATGAGCTGATGCTGCTTGAAGCAATTTAAACATAGATGATGCTGCTATTCAGTA
Db	921 CCAATGAGCTGATGCTGCTTGAAGCAATTTAAACATAGATGATGCTGCTATTCAGTA
Qy	1290 CAGAGATCAACTAGACTTTGAAATATCGACGGCAATTTGCAATTTTGAAGATGAA
Db	981 CAGAGATCAACTAGACTTTGAAATATCGACGGCAATTTGCAATTTTGAAGATGAA
Qy	1350 GGATGGTGTACCAAGTGGAGCATATAAGGAAATAGTATTTTCCGGTACCAACGTCAG
Db	1041 GGATGGTGTACCAAGTGGAGCATATAAGGAAATAGTATTTTCCGGTACCAACATCTAG

Qy	1410 ACGTGATTCCTTGTGGCCATGATTCGCTTCAACAACTCGGAATTTGAAGATATCTTAACA
Db	1101 ACGTGATTCCTTGTGGCCATGATTCGCTTCAACAACTCGGAATTTGAAGATATCTTAACA
Qy	1470 AAGATATGATGTCAGAGCCGGGCAAAATTTTGTACTTATTTGGTGGATGATCGATCGAGC
Db	1161 AAGATATGATTTTGGAGCTGAGCAAAATTTTAGACTTATTTTGGTGGATGATCGATCGAA
Qy	1530 TGACACTAAACCATGATTTT-----ACCAGTTTACATACAAAGCTTTTAAATGTTATACGG
Db	1218 AGAGCTGACAGAAAGTATGACTACCAAGTAGCTACATGCAACATTTTAAATGTTATGGA
Qy	1583 AGAGCTGACTGTCTTAGTGTGAAGGGATATCGGCTTCTTAGTATTTGGAATGATCATCA
Db	1278 AGGAACCCACTGCT-----TATTGTTGGAATGGATGAATCATCA
Qy	1643 ACACAACTATATTTTAAAGTTCAGAACATAAAGAGGAATAGTAGCCCTGTTAAAGACT
Db	1317 CCACATCCTATTA-TTCAAGTTTTTACAAACATAAAGAGGAATAGTAGCCCTTATAACACA
Qy	1703 ATACATGGGACCATCAATATCGCGCGCGCAATT
Db	1376 ATGTTTTGTTTCTAAGAGGAACGTTACGATTT

RESULT 15	1625 bp	mRNA	linear	PLN 19-SEP-2002
BT000334	Arabidopsis thaliana	glycosyltransferase-like protein	(At4g38240)	
LOCUS	mRNA, complete cds.			
ACCESSION	BT000334.1	GI:23198251		
VERSION	BT000334			
KEYWORDS	FLI CDNA			
SOURCE	Arabidopsis thaliana (thale cress)			
ORGANISM	Arabidopsis thaliana			
REFERENCE	1 (bases 1 to 1625)			
AUTHORS	Nguyen,M., Karlin-Neumann,G., Southwick,A., Tripp,M., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kameya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.			
TITLE	Direct Submission			
JOURNAL	Submitted (19-SEP-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA			
COMMENT	e-mail for correspondence: arab@sequence.stanford.edu			

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of AFL CDNAs (RAFL CDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kameya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the pENTR (ORF) clones using the RAFL cDNAs: Nguyen,M., Southwick,A., Tripp,M., Palm,C.J., Jones,T., Wu,T., Chen,H., Cheuk,R., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hsuan,V.W., Lee,J.M., Kim,C.J., Quach,H.L., Shinn,P., Tang,C.C., Totsuami,M., Wallender,E.K., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Ecker,J., Theologis,A. and Davis,R.W.

Nguyen,M., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

location/Qualifiers
1. 1625
/organism="Arabidopsis thaliana"
/mol_type="mRNA"

gene	/db_xref="taxon:3702"			
	/chromosome="4"			
	/clone="U25541"			
	/ecotype="Columbia"			
	/note="This clone is in pENTR/SD-dTopo This is a cloned PCR product using RIKEN clone RAFL09-98-A11 (AY099838) as a template"			
	1. .1625			
	/gene="At4g38240"			
	/note="synonym: F20D10.360"			
	1. .1335			
	CDS	/gene="At4g38240"		
/codon_start=1				
/product="glycosyltransferase-like protein"				
/protein_id="AA15653.1"				
/db_xref="GI:23198252"				
/translation="MARISCDRLFLIPAAFMFIYIQMLFQTSQVADRLSSAIBSE NHCTSMRGILIDEVSIKSRIVALEDMKQDELVOLKDIOTFEKKGIKLSYNOL MPYAAVVMACSRADYKLVKTYQTPVASKYPLFISQDSDQAKVSKLSYNOL TYMHLDFPVPVIERPELTAYYKIAHRYKWLDPYKHKFSRVILEDDMEIADPF FDFEAAASLMRDKTIMAASSWMDNGKQFVHDPYALYKRSDFPGLGMLKSTWDE LSPKWPAYDMDLRLKENHKGQFIRPEVCRTYNFGEHSSIGSQFQSYLEPIKLN DVTWKADLGLYLEGNTYKFSGLVRQARPIQSGDLVLAQNIKDVRIRYKQDVEF ERIAGEFGIPEEWKDGVPRTAYKGVVVFRIQTTRRVFLVGPDSVMQLGRNS"				
ORIGIN				
Query Match 44.4%; Score 772; DB 15; Length 1625;				
Best Local Similarity 73.3%; Pred. No. 4.7e-178;				
Matches 988; Conservative 0; Mismatches 360; Indels 0; Gaps 0;				
QY	139	AAGTTTTGCTGATTTCCGGTACTCTCTCATCTTGGCTGCTGCGCTTCATCTACACA	198	
DB	7	AGGATCTGCTGACITGAGATTTCTTCTCATCCGGCAGCTTTCATGTTCACTACATC	66	
QY	199	CAGATCGGCTTTTGGCACACAGTCAGAAATATGCAGATCGCTTGCTGCTGCAATTGAA	258	
DB	67	CAGATGAGCTTTTCCAGACGCATCACAGTATGCAGATCGCTCAGTTCGGCTATCGAA	126	
QY	259	GCAGAAATCATTTGTACAGCCAGACAGATTTGCTTTATTCACCAGATTAGCCTGCAGCAA	318	
DB	127	TCTGAGAACCAATTGCACTAGTCAAATCGAGGCGCTCATAGATGAAGTTAGCATCAAAACAG	186	
QY	319	GGAGAAATAGTGTCTTTGAAGAACAAATGAAGCGTCAGGACCGAGGAGTGCCGCAATTA	378	
DB	187	TCGGGATTTGTGGCCCTCGAAGATATGAAGAACCGCCAGACGAGAACTTTGTGACGCTT	246	
QY	379	AGGGCTCTGTTT CAGGATCTTGAAGTAAAGGCAATAAAAGTTGATCGGAAATGTACAG	438	
DB	247	AAGGATCTAATCCAGACGTTTGAATAAAAGGAATAGCAAACTCACTCAAGGTGGACAG	306	
QY	439	ATGCCAGTGGCTGCTGTAGTTGTTATGGCTTTGCAATCGGGCTGATTTACCTTGGAAAGACT	498	
DB	307	ATGCCCTGTGGCTGCTGTAGTGGTTATGGCTTGCAGTCTGTGCAGACTATCTTGAAGGAGCT	366	
QY	499	ATTAAATCCATCTTAAATACCAAAATATCTGTTCGCTCAAAATATCTCTTTTCATATCC	558	
DB	367	GTTAAATCAGTTTAAACATATCAAACTCCCGTTGCTTCAAAATATCTCTTATTTATCT	426	
QY	559	CAGGATGGATCACATCTCTGATGT CAGGAAGCTTGCTTTTGAGCTATGATCAGCTGACGTAT	618	
DB	427	CAGGATGGATCTGATCAAGCTGTCAAGAGCAAGTCATTGAGCTATAATCAATTAACATAT	486	
QY	619	ATGCAGCACTTGGATTTTGAACCTGTGCACTACTGAAAGACACAGGGGAGCTGATTGCATAC	678	
DB	487	ATGCAGCACTTGGATTTTGAACCAAGTGGTCACTTGAAGGGCTTGGCGAACTGACTGCGCTAC	546	
QY	679	TACAAATTCGACGTCAATTAAGTGGGCATTGGATCAGCTGTTTTCACAGCAATAATTTT	738	
DB	547	TACAAGATTGACGTCACCTACAAGTGGGCACTGGACCAAGTTGTTTTACAACACAAATTT	606	
QY	739	AGCCGTGTTATCATCTAGAAGATGATATGGAATTCGCCCTGATTTTTTTTGTGACTTTTTT	798	
DB	607	AGTCGAGTGTATTACTAGAAGATGATATGGAATTTGCTCCAGACTTCTTTGATTACTTT	666	

Search completed: December 14, 2005, 11:16:02
Job time : 8633 secs

Qy	799	GAGGCTGGAGTACTCTTCTTGACAGAGCAAGTCGATTATGGCTATTTCTTCTTGAAT 858
Db	667	GAGGCTGCAGCTAGTCTCATGGATAGGATAAAACATTATGGCTGCTTTCATCATGAAT 726
Qy	859	GACAATGGACAATGCAAGTTTGTCCAAAGATCCCTTATGTCTTTTACCGCTCAGATTTTTTT 918
Db	727	GATATGGACAGNAGCAGTTTGTGCATGATCCCTATGCGCTATACCGATCAGATTTTTTT 786
Qy	919	CCCGTCTTTGGATGGATGCTTTCAAAATCTTACTTGGGACGAATATCTCCAAAGTGGCCA 978
Db	787	CTTGGCTTTGGGTGGATGCTCAAGAGATCGACTTGGGATGAGTTATCACCAAGTGGCCA 846
Qy	979	AAGGCTTACTGGGACCACTGGCTAAGACTCAAGAGAAATCACAGAGTGCAGCAATTTAT 1038
Db	847	AAGGCTTACTGGGATGATTTGGCTGAGACTAAAGGAAACATAAAGGCCGCCAATTCAT 906
Qy	1039	CGCCAGAAAGTTTGCAGAAACATATAATTTTGGTGACATGGTTCTTAGTTTGGGGCAGTTT 1098
Db	907	CGACCGGAAGTCTGTAGAACATACATTTTGGTGAACATGGGTCTAGTTTGGGACAGTTT 966
Qy	1099	TTCAAGCAGTATCTTGAGCCAAATTAACCTAAATGTATCCAGTTGATTTGGAAGTCAATG 1158
Db	967	TTCACTCAGTATCTGGAACCTATAAAGCTAAACGATGTGACGTTGACTGGAAGCAAG 1026
Qy	1159	GACCTTAGTTACCTTTTGGAGGACATTAAGTGAACACATTTTGGTGACTTTGGTTAAAAAG 1218
Db	1027	GACCTGGGATACCTTGACAGAGGGAAACTATACCAAGTACTTTTCTGGCTTAGTGAGACAA 1086
Qy	1219	GCTAAGCCCATCCATGGAGCTGATGCTGCTTTGAAAGCATTTTAAACATAGATGGTGTG 1278
Db	1087	GCAGGACCAATTCAGGTTCTGACCTTGCTTAAAGGCTCAAAACATTAAGGATGATGTT 1146
Qy	1279	CGTATTCACTACAGAGATCAACTAGACTTTGAAAAATATCGCACGCAATTTGGCATTTTT 1338
Db	1147	CGTATCCGGTATAAAGACCAAGTAGAGTTTGAACGCAATTGACGGGAAATTTGGTATATTT 1206
Qy	1339	GAAGATGGAAGGATGGTGTACCACTGCGAGCATATAAAGGAATAGTAGTTTCCGGTAC 1398
Db	1207	GAAGAAATGGAAGGATGGTGTGCTCGAACAGCATATAAAGGAGTAGTGGTGTTCGGAATC 1266
Qy	1399	CAAACTGTCAGAGCTGTATTCTTGTGTCGCTTCAACAACTCCGGAATTTGAA 1458
Db	1267	CAGACAAACAGAGCTGTATTCTGTTGGTGGCCAGATTTCTGTAATGACGCTTGGAAATCGA 1326
Qy	1459	GATACCTTAAACAAAGATATGATTCAGGA 1486
Db	1327	AATTCCTGATGCAAAACATATGAAGGA 1354

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 08:25:25 ; Search time 339 Seconds

(without alignments)
9108.037 Million cell updates/sec

Title: US-10-615-144-3

Perfect score: 1737

Sequence: 1 gaattcgccgcccattga.....ataatcgccgcccgaattc 1737

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:**

- 1: /cgn2_6/ptodata/1/ina/1 COMB.seq:**
- 2: /cgn2_6/ptodata/1/ina/5 COMB.seq:**
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:**
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:**
- 5: /cgn2_6/ptodata/1/ina/6 COMB.seq:**
- 6: /cgn2_6/ptodata/1/ina/6CTUS COMB.seq:**
- 7: /cgn2_6/ptodata/1/ina/PP COMB.seq:**
- 8: /cgn2_6/ptodata/1/ina/RE COMB.seq:**
- 9: /cgn2_6/ptodata/1/ina/backfiles1.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1737	100.0	1737	US-09-591-466C-3	Sequence 3, Appl
2	1255	72.3	1669	US-09-591-466C-1	Sequence 1, Appl
3	767.2	44.2	1854	US-09-591-466C-5	Sequence 5, Appl
4	45.4	2.6	1141	US-09-806-708B-22	Sequence 22, Appl
5	45.4	2.6	1141	US-09-806-708B-22	Sequence 22, Appl
6	40	2.3	778	US-09-270-767-9920	Sequence 9920, Ap
7	38.8	2.2	2793	US-09-134-001C-2477	Sequence 2477, Ap
8	38.8	2.2	2976	US-09-386-962C-7	Sequence 7, Appl
9	38.8	2.2	2976	US-09-386-959-7	Sequence 7, Appl
10	38.2	2.2	942	US-09-248-796A-2786	Sequence 2786, Ap
11	36.8	2.1	32172	US-09-949-016-16440	Sequence 16440, A
12	36.6	2.1	2253	US-09-620-312D-194	Sequence 194, Appl
13	36.6	2.1	2725	US-10-012-231A-87	Sequence 87, Appl
14	36.6	2.1	2725	US-10-015-389A-87	Sequence 87, Appl
15	36.6	2.1	2725	US-10-006-768A-87	Sequence 87, Appl
16	36.6	2.1	2725	US-10-015-671A-87	Sequence 87, Appl
17	36.6	2.1	2725	US-10-015-393A-87	Sequence 87, Appl
18	36.6	2.1	2725	US-10-011-833A-87	Sequence 87, Appl
19	36.6	2.1	2725	US-10-006-041A-87	Sequence 87, Appl
20	36.6	2.1	2725	US-10-012-064A-87	Sequence 25, Appl
21	36.6	2.1	2731	US-09-786-240-25	Sequence 25, Appl
22	36.4	2.1	154626	US-09-949-016-14000	Sequence 14000, A
23	36	2.1	858	US-09-107-532A-313	Sequence 313, Appl
24	36	2.1	234884	US-09-949-016-16420	Sequence 16420, A

Sequence 11259, A
Sequence 16, Appl
Sequence 16, Appl
Sequence 16101, A
Sequence 2813, Ap
Sequence 49, Appl
Sequence 12674, A
Sequence 13724, A
Sequence 13724, A
Sequence 13, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 12584, A
Sequence 35, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 854, App
Sequence 12505, A
Sequence 14207, A
Sequence 40, Appl

ALIGNMENTS

RESULT 1

US-09-591-466C-3
; Sequence 3, Application US/09591466C
; Patent No. 6653459
; GENERAL INFORMATION:
; APPLICANT: Von Schaeuwen, Antje
; TITLE OF INVENTION: Plant GntI Sequences and the Use Thereof for the Production
; of Plants Having Reduced or Lacking N-acetyl Glucosaminyl
; TITLE OF INVENTION: Transfrase I(GntI) Activity
; FILE REFERENCE: 032266-003
; CURRENT APPLICATION NUMBER: US/09/591.466C
; CURRENT FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: EP 98/08001
; PRIOR FILING DATE: 1998-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (733)...(741)
; OTHER INFORMATION: function: Asn codon in this context is a potential
; glycosylation site;
; OTHER INFORMATION: glycosylation site;
; OTHER INFORMATION: product: N-glycosylation consensus sequence;
; OTHER INFORMATION: phenotype: N-glycans modulate protein properties;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (733)...(741)
; OTHER INFORMATION: standard name: N-glycosylation site;
; OTHER INFORMATION: label: pgt-CHO;
; OTHER INFORMATION: note: gntI sequences from animals do not contain
; this feature.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (127)...(1467)
; OTHER INFORMATION: codon_start: 127;
; OTHER INFORMATION: function: initiates complex N-glycans on secretory
; glycoproteins;
; OTHER INFORMATION: glycoproteins;
; OTHER INFORMATION: EC_number: 2.4.1.101;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (127)...(1467)
; OTHER INFORMATION: product: beta-1,2-N-acetylglucosaminyltransferase I;
; OTHER INFORMATION: evidence: EXPERIMENTAL;
; OTHER INFORMATION: gene: cgl;

= parent of restant
appl

```

; OTHER INFORMATION: standard_name: gntI;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (127)...(1467)
; OTHER INFORMATION: label: ORF;
; OTHER INFORMATION: note: first gntI sequence from tobacco (unpublished).
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (15)...(126)
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (1468)...(1723)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (154)...(213)
; OTHER INFORMATION: function: membrane anchor (amino acids 10-29);
; OTHER INFORMATION: product: hydrophobic amino acid stretch in GntI;
; OTHER INFORMATION: standard_name: membrane anchor of a type II golgi;
; OTHER INFORMATION: protein.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(14)
; OTHER INFORMATION: function: use for cloning the cDNA library in
; OTHER INFORMATION: Lambda ZAPII;
; OTHER INFORMATION: product: EcoRI/No. 6653459I-cDNA adapter;
; OTHER INFORMATION: number: 1.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1724)...(1737)
; OTHER INFORMATION: product: EcoRI/No. 6653459I-cDNA adapter;
; OTHER INFORMATION: number: 2.
; US-591-466C-3

Query Match      100.0%; Score 1737; DB 3; Length 1737;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1737; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1  GNAATCGGGCCGCGCATGCTGATCCTAACTGAAAGGCAAGGCAAGTAAATCCAGCGATGA 60
Db 1  GNAATCGGGCCGCGCATGCTGATCCTAACTGAAAGGCAAGTAAATCCAGCGATGA 60

Qy 61  AACACTCATAACTGAAACACTGAGAGACTATTTCGCTTCTCTAAAGCCTTCAATCGAATT 120
Db 61  AACACTCATAACTGAAACACTGAGAGACTATTTCGCTTCTCTAAAGCCTTCAATCGAATT 120

Qy 121  CGCACGATGAGAGGAAACAAAGTTTGTGCTGTAATTCGCGTACCTCTCATCTTGGCTGCT 180
Db 121  CGCACGATGAGAGGAAACAAAGTTTGTGCTGTAATTCGCGTACCTCTCATCTTGGCTGCT 180

Qy 181  GTCCCTTCATCTACACACAGATCGCGCTTTTTCGCGACACAGTCAGAAATATGACATCGC 240
Db 181  GTCCCTTCATCTACACACAGATCGCGCTTTTTCGCGACACAGTCAGAAATATGACATCGC 240

Qy 241  CTTGCTGCTGCAATTTGAAGCAGAAATCATTTGTACAAAGCCAGACAGATTGCTTATTGAC 300
Db 241  CTTGCTGCTGCAATTTGAAGCAGAAATCATTTGTACAAAGCCAGACAGATTGCTTATTGAC 300

Qy 301  CAGATTAGCCTGACGAAAGGAAGTAATGTTGCTCTTGAAGAACAAATTAAGACGCTCAGGAC 360
Db 301  CAGATTAGCCTGACGAAAGGAAGTAATGTTGCTCTTGAAGAACAAATTAAGACGCTCAGGAC 360

Qy 361  CAGAGTGGCCGACAAATTAAGGCTCTTGTTCAGGATCTTGAAGTAAAGGCGATAAAAAAG 420
Db 361  CAGAGTGGCCGACAAATTAAGGCTCTTGTTCAGGATCTTGAAGTAAAGGCGATAAAAAAG 420

Qy 421  TTGATCGGAAATGTACAGATGCCAGTGGCTGCTGTAGTTGTATGCTTCAATTCGGCT 480
Db 421  TTGATCGGAAATGTACAGATGCCAGTGGCTGCTGTAGTTGTATGCTTCAATTCGGCT 480

Qy 481  GATTACCTGGAAGAAGCTATTAAATCCATCTTAAATACCAAAATATCTGTTGGCTCAAAA 540
Db 481  GATTACCTGGAAGAAGCTATTAAATCCATCTTAAATACCAAAATATCTGTTGGCTCAAAA 540

; OTHER INFORMATION: standard_name: gntI;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (127)...(1467)
; OTHER INFORMATION: label: ORF;
; OTHER INFORMATION: note: first gntI sequence from tobacco (unpublished).
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (15)...(126)
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (1468)...(1723)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (154)...(213)
; OTHER INFORMATION: function: membrane anchor (amino acids 10-29);
; OTHER INFORMATION: product: hydrophobic amino acid stretch in GntI;
; OTHER INFORMATION: standard_name: membrane anchor of a type II golgi;
; OTHER INFORMATION: protein.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(14)
; OTHER INFORMATION: function: use for cloning the cDNA library in
; OTHER INFORMATION: Lambda ZAPII;
; OTHER INFORMATION: product: EcoRI/No. 6653459I-cDNA adapter;
; OTHER INFORMATION: number: 1.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1724)...(1737)
; OTHER INFORMATION: product: EcoRI/No. 6653459I-cDNA adapter;
; OTHER INFORMATION: number: 2.
; US-591-466C-3

541  TATCCTCTTTTTCATATCCAGGATGGATCACATCCTGATGTCAGGAAGCTTGTCTTTCAGC 600
541  TATCCTCTTTTTCATATCCAGGATGGATCACATCCTGATGTCAGGAAGCTTGTCTTTCAGC 600
601  TATGATCAGCTGACGTATATGACAGCACTTGGATTTTGAACCTGTGCATATCTGAAAGACCA 660
601  TATGATCAGCTGACGTATATGACAGCACTTGGATTTTGAACCTGTGCATATCTGAAAGACCA 660
661  GGGGAGCTGATTCGATACATACAAAAATTGACAGTCAATTAAGAATGGGATTCGATTCAGCTG 720
661  GGGGAGCTGATTCGATACATACAAAAATTGACAGTCAATTAAGAATGGGATTCGATTCAGCTG 720
721  TTTTCAAGCATAAATTTTAGCCGTGTATATCATATACTAAGATGATATGGAATTTGCCCT 780
721  TTTTCAAGCATAAATTTTAGCCGTGTATATCATATACTAAGATGATATGGAATTTGCCCT 780
781  GATTTTTCATCTTTTTCAGGCTGGAGCTACTCTTCTTGCACAGACAAAGTGCATTTATG 840
781  GATTTTTCATCTTTTTCAGGCTGGAGCTACTCTTCTTGCACAGACAAAGTGCATTTATG 840
841  GCTATTTCTTTTTCAGGATGACAAATGACAAATGCAATGCTTTCGCAAGATCCTTATGCTCTT 900
841  GCTATTTCTTTTTCAGGATGACAAATGACAAATGCAATGCTTTCGCAAGATCCTTATGCTCTT 900
901  TACCGCTCAGATTTTTCGCGCTTTCGATGGATGCTTTCGCAAGTCTTCAAAATCTACTTGGACGAA 960
901  TACCGCTCAGATTTTTCGCGCTTTCGATGGATGCTTTCGCAAGTCTTCAAAATCTACTTGGACGAA 960
961  TTATCTCCAAAGTGGCCAAAGGCTTACTTGGGACGACTGGCTAAGACTCAAGAGAAATCAC 1020
961  TTATCTCCAAAGTGGCCAAAGGCTTACTTGGGACGACTGGCTAAGACTCAAGAGAAATCAC 1020
1021  AGAGTGCAGCAATTTATTCGCCAGAAATTTTCAGAAATGCAATATATTTTGTGTGAGCATGGT 1080
1021  AGAGTGCAGCAATTTATTCGCCAGAAATTTTCAGAAATGCAATATATTTTGTGTGAGCATGGT 1080
1081  TCTAGTTTCGGGCGAGTTTTCGAGCAGTATCTTGAAGCAATTAATTAATTAATTAATTAATTAAT 1140
1081  TCTAGTTTCGGGCGAGTTTTCGAGCAGTATCTTGAAGCAATTAATTAATTAATTAATTAATTAAT 1140
1141  GTTGATTTGGAAGTCAATGGAACCTTACTTGTACCTTTTGGAGGACAAATTAACGTAACACTTTT 1200
1141  GTTGATTTGGAAGTCAATGGAACCTTACTTGTACCTTTTGGAGGACAAATTAACGTAACACTTTT 1200
1201  GGTGACTTGGTTTAAAGGCTAAGCCATTCATGAGAGTGTATGCTGCTTGTGAAAGCATTT 1260
1201  GGTGACTTGGTTTAAAGGCTAAGCCATTCATGAGAGTGTATGCTGCTTGTGAAAGCATTT 1260
1261  AACATAGATGGTATGCTGCTATTCAGTACAGAGATCACTAGACTTTGAAATATATCGCA 1320
1261  AACATAGATGGTATGCTGCTATTCAGTACAGAGATCACTAGACTTTGAAATATATCGCA 1320
1321  CGGCAATTTGGCAATTTTGAAGAAATGGAAGGATGGTGTACACGTCAGCATATAAAGGA 1380
1321  CGGCAATTTGGCAATTTTGAAGAAATGGAAGGATGGTGTACACGTCAGCATATAAAGGA 1380
1381  ATAGTAGTTTTCGGTACCAACGTCAGAGCTGTATTCCTTGTGCGCATGATTCGCTT 1440
1381  ATAGTAGTTTTCGGTACCAACGTCAGAGCTGTATTCCTTGTGCGCATGATTCGCTT 1440
1441  CAACTAGCTGGAATTCAGATACCTTAACAAAGATATGATTCAGGAGGAGCCCGGCAAAAT 1500
1441  CAACTAGCTGGAATTCAGATACCTTAACAAAGATATGATTCAGGAGGAGCCCGGCAAAAT 1500
1501  TTTGACTTATTTGGGTAGGATGCAATCGAGCTGACACTAAACCAATGATTTTACCAGTTACAT 1560
1501  TTTGACTTATTTGGGTAGGATGCAATCGAGCTGACACTAAACCAATGATTTTACCAGTTACAT 1560
1561  ACAACGTTTAAATGTTATACGAGGAGCTCACTGTTCTAGTGTGTAAGGGATATTCGCTT 1620
1561  ACAACGTTTAAATGTTATACGAGGAGCTCACTGTTCTAGTGTGTAAGGGATATTCGCTT 1620
1621  CTTAGTATTGGATGAATCATCAACCAACCTATTATTTTAAAGTGTTCAGAAATATAAAGAG 1680
```


Db 1621 CTTAGTATTGATGATCATCAACACACCTATTATTATTTAAGTGTTCAGAACATAAAGAG 1680
 Qy 1681 GAAATGAGCCCTGTAAGACTATACATGAGACCATCATATATCCGCGCGGAATTC 1737
 Db 1681 GAAATGAGCCCTGTAAGACTATACATGAGACCATCATATATCCGCGCGGAATTC 1737

RESULT 2
 US-09-591-466C-1
 ; Sequence 1, Application US/09591466C
 ; Patent No. 6653459
 ; GENERAL INFORMATION:
 ; APPLICANT: Von Schaeven, Antje
 ; TITLE OF INVENTION: Plant GntI Sequences and the Use Thereof for the Production
 ; TITLE OF INVENTION: of Plants Having Reduced or Lacking N-acetyl Glucosaminyl
 ; TITLE OF INVENTION: Transferase I (GntI) Activity
 ; FILE REFERENCE: 032266-003
 ; CURRENT APPLICATION NUMBER: US/09/591,466C
 ; CURRENT FILING DATE: 2000-06-09
 ; PRIOR APPLICATION NUMBER: EP 98/08001
 ; PRIOR FILING DATE: 1998-09-12
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 1669

; TYPE: DNA
 ; ORGANISM: Solanum tuberosum
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (659)...(667)
 ; OTHER INFORMATION: function: Asn codon in this context is a potential
 ; OTHER INFORMATION: glycosylation site;
 ; OTHER INFORMATION: product: N-glycosylation consensus sequence;
 ; OTHER INFORMATION: phenotype: N-glycans modulate protein properties;
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (659)...(667)
 ; OTHER INFORMATION: standard_name: N-glycosylation site;
 ; OTHER INFORMATION: label: pot-CHO;
 ; OTHER INFORMATION: note: GntI-coding sequences from animals do not
 ; OTHER INFORMATION: contain this feature.
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (53)...(1393)
 ; OTHER INFORMATION: codon_start: 53;
 ; OTHER INFORMATION: function: initiates complex N-glycans on secretory
 ; OTHER INFORMATION: glycoproteins;
 ; OTHER INFORMATION: EC_number: 2.4.1.101;
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (53)...(1393)
 ; OTHER INFORMATION: product: beta-1,2-N-acetylglucosaminyltransferase I;
 ; OTHER INFORMATION: evidence: EXPERIMENTAL;
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (53)...(1393)
 ; OTHER INFORMATION: gene: cgl;
 ; OTHER INFORMATION: standard_name: gntI;
 ; OTHER INFORMATION: label: ORF;
 ; OTHER INFORMATION: note: first gntI sequence from potato (unpublished).
 ; FEATURE:
 ; NAME/KEY: 5'UTR
 ; LOCATION: (15)...(52)
 ; FEATURE:
 ; NAME/KEY: 3'UTR
 ; LOCATION: (1394)...(1655)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (80)...(139)
 ; OTHER INFORMATION: function: membrane anchor (amino acids 10-29);
 ; OTHER INFORMATION: product: hydrophobic amino acid stretch in GntI;
 ; OTHER INFORMATION: standard_name: membrane anchor of a type II Golgi

; OTHER INFORMATION: protein;
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (80)...(139)
 ; OTHER INFORMATION: note: identified by comparison with GntI sequences
 ; OTHER INFORMATION: from animals.
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)...(14)
 ; OTHER INFORMATION: function: used for cloning the cDNA library in
 ; OTHER INFORMATION: Lambda ZAPII;
 ; OTHER INFORMATION: product: EcoRI/No. 66534591-cDNA adapter;
 ; OTHER INFORMATION: number: 1.
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1655)...(1669)
 ; OTHER INFORMATION: product: EcoRI/No. 66534591-cDNA adapter;
 ; OTHER INFORMATION: number: 2.
 ; US-09-591-466C-1
 Query Match 72.3%; Score 1255; DB 3; Length 1669;
 Best Local Similarity 89.1%; Pred. No. 0;
 Matches 1412; Conservative 0; Mismatches 140; Indels 33; Gaps 4;
 Qy 126 GATGAGAGGGAACAAAGTTTGGCTGTGATTTCCGTAACCTCTCATCTCTGCTGTGCTGCTGC 185
 Db 52 GATGAGAGGGAACAAAGTTTGGCTGTGATTTACCGTACCTCTCTCGTGGCTGTGCTGCTGC 111
 Qy 186 CTTTCATCTACACAGATGCGGCTTTTGGACACAGTCAGAAATATGAGATCGCTTGC 245
 Db 112 CTTTCATCTACACAGATGCGGCTTTTGGACACAGTCAGAAATATGAGATCGCTTGC 171
 Qy 246 TGCTGCAATTGAAGCAGAAATCATTTGACAAAGCCAGACAGATGCTTTATTGACAGAT 305
 Db 172 TGCTGCAATTGAAGCAGAAATCATTTGACAAAGCCAGACAGATGCTTTATTGACAGAT 231
 Qy 306 TAGCTCTGACGACGAAAGAAATAGTTGCTCTTGAAGAACAAATGAAGCTCAGGACGAGA 365
 Db 232 TAGCCAGCAGCAGGAGAGTAGTAGCTCTTGAAGAACAAATGAAGCATCAGGACGAGA 291
 Qy 366 GTGCCGCAATTAAGGCTCTTGTTCAGGATCTTGAAGTAAGGCATAAAGATTCAT 425
 Db 292 GTGCCGCAATTAAGGCTCTTGTTCAGGATCTTGAAGTAAGGCATAAAGATTCAT 351
 Qy 426 CGGAATCTACAGATGCCAGTGGCTGTGATGTTGTTATGCTGCAATCGGCTGATTA 485
 Db 352 CGGAGATGTGAGATGCCAGTGGCAGCTGTGATGTTGTTATGCTGCAATCGGCTGATTA 411
 Qy 486 CCTGGAAGACTATTAAATCCATCTTAAATAACAAATATCTGTTGCTCAAAATATCC 545
 Db 412 CTTGAGAGGACTATTAAATCCATCTTAAATAACAAATATCTGTTGCTCAAAATATCC 471
 Qy 546 TCTTTTCATATCCAGGATGATCATCTCTGATGTAGGAGCTTGTCTTTCAGCTATGA 605
 Db 472 TCTTTTCATATCCAGGATGATCAATCTCTGATGTAGGAGCTTGTCTTTCAGCTATGA 531
 Qy 606 TCAGCTGACGTATATGACGACTTGGATTTTGAACCTGTGCATCTGAAAGACCGGGA 665
 Db 532 TCAGCTGACGTATATGACGACTTGGATTTTGAACCTGTGCATCTGAAAGACCGGGA 591
 Qy 666 GCTGATTCATATACAAATTTGACGCTCATTTACAAGTGGGCAATTTGATCGCTTTTA 725
 Db 592 ACTGTTGATATACAGATTTGACGCTCATTTACAAGTGGGCAATTTGATCGCTTTTA 651
 Qy 726 CAAGCATATTTTAGCCGTGTATCATATCTAGAGATGATATGGAATTTGCCCTGATTT 785
 Db 652 CAAGCATATTTTAGCCGTGTATCATATCTAGAGATGATATGGAATTTGCCCTGATTT 711
 Qy 786 TTTTTCATTTTGGGCTGGAGCTTCTCTTTCACAGAGACAAAGTCGATTTATGCTAT 845
 Db 712 TTTTTCATTTTGGGCTGGAGCTTCTCTTTCACAGAGACAAAGTCGATTTATGCTAT 771
 Qy 846 TTTTTCATTTTGGGCTGGAGCTTCTCTTTCACAGAGACAAAGTCGATTTATGCTAT 905

[illegible]

RESULT 3

RESULTS 3
US-09-591-466C-5

US-09-591-466C-5
: Semience 5. Application IIS/09591466C

; Sequence 5, Application No. 6653459

Patent No. 6653459

; GENERAL INFORMATION:

APPLICANT: Von Schaewen, Antje

; TITLE OF INVENTION: Plant GntI Sequences and the Use Thereof for the Production

; TITLE OF INVENTION: of Plants Having Reduced or I

	OF INVENTION:	OF FI
; TITLE OF INVENTION:	Trans	
; TITLE OF INVENTION:	Trans	

FILE OF INVENTION: TRANSFERASE I (GIII)
FILE REFERENCE: 032266-003

FILE REFERENCE: 032266-
CURRENT APPLICATION NUMBER:

; CURRENT APPLICATION NUMBER: US/09/591
 : CURRENT FILING DATE: 2000-06-09

Query Match 44.2%: Score 767.2: DB 3: Length 1854:

Query Match	44.2%	score	167.2	DB 3
Best [local] Similarity	73.1%	Pred	NO	1 6e-232

Best Local Similarity 73.1%; Pred: No. 1.6e-232;
Matches 985; Consistent 0; Mismatches 252;
Indels 0; Cons 0

139 AAGTTTGGTGTGATTTCGGGTACCTTCCTCATCTTGGCTGCTGTGGCTTCATCTACACA 198

139 AAGTTTGGCTGTGATTCCGGTACCTCCCTCATCTTGGCTGCTGCTGCCTTCATCTACACA 198

PRIOR FILING DATE: 1998-09-12
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 5
LENGTH: 1854
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1185)...(1193)
OTHER INFORMATION: function: Asn Codon is a potential glycosylation site;
OTHER INFORMATION: product: Consensus sequence for N-glycosylation;
OTHER INFORMATION: phenotype: N glycans modulate protein characteristics;
OTHER INFORMATION: standard name: N glycosylation site;
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1185)...(1193)
OTHER INFORMATION: label: pot-CHO;
OTHER INFORMATION: note: absent in animal GnTI sequences.
FEATURE:
NAME/KEY: CDS
LOCATION: (135)...(1469)
OTHER INFORMATION: codon_start: 135;
OTHER INFORMATION: function: initiates complex N glycans on secretory glycoproteins;
OTHER INFORMATION: EC_number: 2.4.1.101;
FEATURE:
NAME/KEY: CDS
LOCATION: (135)...(1469)
OTHER INFORMATION: product: beta-1,2-N-acetyl glucosaminyl transferase I;
OTHER INFORMATION: evidence: EXPERIMENTAL;
OTHER INFORMATION: gene: cgl;
OTHER INFORMATION: standard_name: gntI;
FEATURE:
NAME/KEY: CDS
LOCATION: (135)...(1469)
OTHER INFORMATION: label: ORF;
OTHER INFORMATION: note: first gntI sequence from Arabidopsis
OTHER INFORMATION: (unpublished).

Ov

Qy

Qy	199	CAGATCGGCTTTTTCGCACACAGTCAGATAGCAGATCGCCTTCTGCTGCTCAATTGAA	258
Db	201	CAGATAGGCTTTTTCGACAGCAATCACAGTATGTCAGATCGCCTCAGTTCCGCTATPCGAA	260
Qy	259	GCAGAAATCAATTGTATCAACGCCAGACAGATTCCTTATTGACCAGATTAAGCTCGCAGCAA	318
Db	261	TCTGAGAACCAATTGCATAGTCAAAATCGGAGGCTCATAGATGAAGTTAGCATCAACAG	320
Qy	319	GGAGAAATAGTTGCTCTTTGAAGAACAAATGAAGCGTTCAGGACCGAGAGTGCAGCAAA	378
Db	321	TCGCGGATTGTGCGCTTCGAAGATATCAAGAAACCGCCAGGACGAAGAACTTGTGCGAGCTT	380
Qy	379	AGGCTCTTGTTCAGGATCTTTGAAGCTTAAGGGCATAAABAAAGTTGATCGGAAATGTACAG	438
Db	381	AAGGATCTAAATCCAGACGTTTGAABAAAAAAGSAAATAGCAAAATCTCACTCAAGGTGCACAG	440
Qy	439	ATGCCAGTGGCTGCTGTAGTTGTTATGGCTTGCAATCGGGCTGATTAACCTTGGAAAAAGACT	498
Db	441	ATGCCGTGTGCTGCTGTAGTTGTTATGGCCCTGCAGTCTGCAGACTATCTTGAAGGACT	500
Qy	499	ATTAATAATCCATCTTTAAATAACCAATAATCTGTTGGCTCAAAATATCTCTTTTCATATCC	558
Db	501	GTTAAATCAGTTTTTAACATATCAAACTCCCGTTGCTTCAAAATATCTCTATTATATATCT	560
Qy	559	CAGATCGATCAATCTGATGATCAGAAAGCTTGTCTTTGAGCTATGATCAGCTGACGTAT	618
Db	561	CAGGATGGATCTGATCAAGCTGTCAAGAGCAAGTCATTGAGCTATTAATCAATTAACATAT	620
Qy	619	ATGCAGACATTGGATTTTGAACCTGTGCATACTGMAAGACACAGGGAGCTGATGTCATAC	678
Db	621	ATGCAGACATTGGATTTTGAACCGTGGTCACTGGAAGGCTCTGGTGNACTGACTGCGTAC	680
Qy	679	TACAAAAATGCAGCTCATTACAAGTGGCCATTGGATCAGCTGTTTTCACAAGCAATAATTT	738
Db	681	TACAAGATTGCACGCTACTCAAGTGGGCACCTGGACCAGTTGTTTTCACAAACACAAATTT	740
Qy	739	AGCCGTGTTATCATACTAGAAAGATGATATGMAAATGCCCCCTGATTTTTTTTGACTTTTTT	798
Db	741	AGTCGAGTGATTATACTAGAAAGCATATGMAAATTGCTCCAGACTTCTTTTGATTACTTTT	800
Qy	799	GAGGCTCGAGCTACTCTTTCTTGACAGAGACAAGTCGATATGGCTATTTCTCTTGGAAT	858
Db	801	GAGGCTGCAGCTAGTCTCATGGATAGGGATAAACCAATTATGGCTGCTTCATCATGGAT	860
Qy	859	GACAAATGGACAAATTCGAGTTTGTCCAAAGATCCTTATGCTCTTTTACCGCTCAGATTTTTT	918
Db	861	GATAAATGGACAGACGAGTTTGTGCTAGTATCCCTATGCGCTATATACCGATCAGATTTTTT	920
Qy	919	CCCGGCTTTGGATGGATGCTTTTCAAAATCTATCTTTGGGACGAAATATCTCCAAAGTGGCCA	978
Db	921	CCTGGCCTTGGGTGGATGCTCAAGAGATCGACTTGGGATGAGTTATCACAAAGTGGCCA	980
Qy	979	AAGGCTTACTGGGACGACTGGCTTAAGACTCAAGAGAAATCACAGAGTCCGACAAATTATTT	1038
Db	981	AAGGCTTACTGGGATGATTTGGCTCAGACTAAAGGAAACCAATAAAGGCCCAATTCATT	1040
Qy	1039	CGCCACAGAAATTTCAGAAACATATAATTTTGGTGAGCATGGTTCTAGTTTGGGCGAGTTT	1098
Db	1041	GCACGGGAAGTCTGTAGNAACATACAAATTTTGGTGNAACATGGGCTCTAGTTTGGGACAGTTT	1100
Qy	1099	TTCAAGCAGTATCTTTGAGCCAAATTAACATAATGATGTCGAGGTGTGATTTGGAAGTCAATG	1158
Db	1101	TTCACTCAGTATCTGGAACCTATAAAGCTAAACGATGTGACGGTTGACTCGAAAGCAAG	1160
Qy	1159	GACCTTAGTTACCTTTTGGAGGACAAATTAACGTGAAACACTTTTGGTGTGACTTGGTTAAAAAG	1218
Db	1161	GACCTGGGATACCTGCACAGAGGGAAACCTATACCAAGTACTTTTCTGGCTTAGTGAGACAA	1220
Qy	1219	GCTAAGCCCATCTCAATGGAGCTGATGCTGCTTTGAAAGCATTTAAACATAGATGGTGAATG	1278
Db	1221	GCACGACCAATTTCAAGGTTTCTGACCTTGTCTTTAAGGCTCAAAACATTAAGGATGATGAT	1280
Qy	1279	CGTATTCAGTACAGAGATCAACTAGACTTTGAAAATATCGCACCGGCAATTTTGGCATTTTT	1338

```

Db      1281  CGTATCCGTTATTAAGACCAAGTAGAGATTGGAACGATTGCAGGGGAATTTGGTATATTT 1340
Qy      1339  GAAGAATGGAAGGATCGGTGTACACATCGTCAGCATATAAAGGAATAGTAGTTTCCGGTAC 1398
Db      1341  GAAGAATGGAAGGATCGGTGTGCCACCAACAGCATATAAAGGAGTAGTGGTGTTCGAATC 1400
Qy      1399  CAAACGTCACAGACGTGTATTTCCTTGTGTGGCCATGATTTCGCTTCAACAACTCGGAATTGAA 1458
Db      1401  CAGACAAACAAGACGTGTATTTCCTGTGTGGCCAGATTCGTAATGACAGCTTGGAAATCGA 1460
Qy      1459  GATACCTTAACAAAGATATGATTCAGGA 1486
Db      1461  AATCTGATGCAAAACATATGAAGA 1488

RESULT 4
US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEL promoter
US-09-806-708B-22

```

Db 465 CTHNNTTWKMKTYNNCKWMSWNGSHRBAABAVTYWYWWRRYAHANNNNNDYWKKA 524
Qy 873 GCAGTTTGTCAAGATCCTTATGCTCTTTACCGCTCAGATTTTTTCCCGTCTTGATG 932
Db 525 CTWYKYBVCWNNYAAWYTKSSWNTSYRYKTKNNRWRSRDTSRMGRANNYARABH 584
Qy 933 GATGCTTTCAAAATCTACTTGGGA--CGAATTATCTCCAAGTGGCCAAAGGCTTACTG 989
Db 585 YGYKWNTRWBWSHTWBHBRAGAAYWYBMYBAKCHMKAWYKAKKYAGAGGSNNNNNN 644
Qy 990 GGACGACTGGCTAAGACTCAAGAGAGATCAAGAGGTTCGACAAATTTATTCGCCCAGAAGT 1049
Db 645 NNNNNNNNATCARDYAAASRYAMANAkWYYKEAANNAYYTHANNWGCWNNATDT 704
Qy 1050 TTGCAGAACATATAATTTTGGTGAGCATGGTCTTAGTTTGGGCGAGTTTTTCAAGCAGTA 1109
Db 705 RRTWKNNNNNNAGTWKNNNNNAKNAKAAKNAKAAKNAKAAKNAKNAKNAKNAKNAKNA 764
Qy 1110 TCTTGAG 1116
Db 765 AABTTDK 771

RESULT 5
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; PRIOR FILING DATE: 2001-04-03
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.c., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

Query Match 2.6%; Score 45.4; DB 3; Length 1141;
Best Local Similarity 10.8%; Pred. No. 0.0018;
Matches 63; Conservative 202; Mismatches 317; Indels 0; Gaps 0;
Qy 764 ATATGGAATTCGCCCTGATTTTGTGACTTTTGTGAGCTGGAGTACTCTTCTGCACA 823
Db 915 WTACYNRAATNNKATWMMKTHGASHKSRTRHTRICRRTKYNNNNNNNARTVYVYHAA 856
Qy 824 GAGACAAGTCAGTATGCTATTTCTTCTTGAATGACAATGGAACAATGCAAGTTTGTCC 883
Db 855 RRRWNAWTRTNNNNNNNNNACRTRTWABWKHSWNNNNNNNNNNNNNNNNNNNNNNNN 796
Qy 884 AAGATCCTTGCTCTTTACCGCTCAGATTTTTCGGGCTCTGGATGGATGCTTCAA 943
Db 795 NABECYRANNNAARWARTCNMYHAAVTTHTDWCYKTMWNTWYDMMTMBTTTTRN 736
Qy 944 AATCTACTTGGGAGCAATATCTCAAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTAA 1003
Db 735 MTSTNNNNNNNNWACTNNNNNNNNKAYYAHATNNGWGNWNTDARRNNNTVMRRRW 676
Qy 1004 GACTCAAGAGAAATCAGAGGTGCAAAATTTATTCGCCCGAAGATTTGCGAGAACATATA 1063
Db 675 MTNTKTRWYSTTRRHVYTGATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 616
Qy 1064 ATTTTGGTGAGCATGGTCTTAGTTTGGGGCAGTTTTTCAAGCAGCATATCTTTCGCAATTA 1123

Db 615 KKVKMRDTTCTYVDVWADSWWYANWMCRDVYTRNNYCKSYAHSYWWSNNAMWYR 556
Qy 1124 AACTAAATGATGTCCAGGTGATTGGAAGTCAATGAGACCTTAGTACCTTTTGGAGACA 1183
Db 555 RYSARNWSSMARWTTTRNNNMMWSGBVRMRWAGTMMWRHNNNNNTDTRYWMMKWARBT 496
Qy 1184 ATTACGTGAACACATTTTGGTGACTTGGTTTAAAGAGCTAAGCCCATCCATGGAGCTGATG 1243
Db 495 VYDSMCNAKSMWRGNWRAKMMWAAANNNDAGAMDHWYTWGNNNTMMRRRAKMMNMAWCR 436
Qy 1244 CTCTCTTTGAAAGCATTTAAACATAGATGCTGATTCGCTATTTCAGTACAGAGATCAACTAG 1303
Db 435 RAVCCNNNNNRACVWHKHKMRWTKYMMKACANNNNBRKAMYRVAWMMYSRDTTNTDWM 376
Qy 1304 ACTTTGAAATATCGCAGCGCAATTTGGCATTTTTTGAAGAAAT 1345
Db 375 MMTSDWBWHWYTVDYTMRAWNNNNNNNNNNNRBCKTTTSMWMMMD 334

RESULT 6
US-09-270-767-9920/c
; Sequence 9920, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9920
; LENGTH: 778
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-9920

Query Match 2.3%; Score 40; DB 3; Length 778;
Best Local Similarity 47.6%; Pred. No. 0.075;
Matches 118; Conservative 0; Mismatches 130; Indels 0; Gaps 0;
Qy 1155 AATGGACCTTAGTACTTTTGGAGGACAATTAAGTGAACACTTTGGTGACTTGGTTAA 1214
Db 756 AATCAATATGAGTACTCTGCTGAAGACAATTAAGTGAACACTTTTTCGCGCGCGTTA 697
Qy 1215 AAAGGCTAAGCCCATCCATGGAGCTGATGCTGTTGAAAGCATTTAAACATAGATGTGA 1274
Db 696 TAGTATCCCATTTGTTACGTACGATGAGCTGCGCGGAACCTGATTAGATCGAAGTCC 637
Qy 1275 TGTGCGTATTCAGTACAGAGATCAACTAGACTTTTGAATAATATCGACGGCAATTTGGCAT 1334
Db 636 AGTTCGATTCATTAATACTACTAGGAGCAGTACAAGCGGACAACTAAGATGCTGGGTCT 577
Qy 1335 TTTTGAAGATGAAGGATGTTACCACTGAGTACAGTATTAAGGATATAGTATTTTCCG 1394
Db 576 TATGGATGATTTCAAGAGCGGTGTTCCCGGAGTGCCTACCATGGCATCGTCTCTCTTA 517
Qy 1395 GTACCAAA 1402
Db 516 CTACAACA 509

RESULT 7
US-09-134-001C-2477
; Sequence 2477, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C

```
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2477
; LENGTH: 2793
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; US-134-001C-2477

Query Match      2.2%; Score 38.8; DB 3; Length 2793;
Best Local Similarity 51.8%; Pred. No. 0.41;
Matches 88; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 1213 AAAAAGGCTAAGCCCATCCATGGAGCTGATGCTCTTTGAAAGCAATTTAACAATAGATGGT 1272
Db 2035 ACAATCTGCTAGACTCAGAAGGCAATTTCTGTATGGGTAACTATTAAACGGACAAGAC 2094

Qy 1273 GATGTGGGTATTTCAGTACAGAGATCAACTAGACTTTGAAATATCGCACGCAATTTGGC 1332
Db 2095 GATATGACTATTGTATAGCGGATTTTATCAACACACCTAAATATAGCTTAGGGAATATGTA 2154

Qy 1333 ATTTTGAAGATGGAAGGATGGTGTACCACTGCGACGATATAAAGGAAT 1382
Db 2155 TGGTATGACACTAATGAAGATGGTATTCAAGGTGATGATGAAGGAAT 2204

RESULT 8
US-09-386-962C-7
; Sequence 7, Application US/09386962C
; Patent No. 6635473
; GENERAL INFORMATION:
; APPLICANT: POSTER, Timothy
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS
; FILE REFERENCE: P06335US2/BAS
; CURRENT APPLICATION NUMBER: US/09/386,962C
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 2976
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(2975)
; US-09-386-962C-7

Query Match      2.2%; Score 38.8; DB 3; Length 2976;
Best Local Similarity 51.8%; Pred. No. 0.41;
Matches 88; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 1213 AAAAAGGCTAAGCCCATCCATGGAGCTGATGCTCTTTGAAAGCAATTTAACAATAGATGGT 1272
Db 2136 ACAATCTGCTAGACTCAGAAGGCAATTTCTGTATGGGTAACTATTAAACGGACAAGAC 2195

Qy 1273 GATGTGGGTATTTCAGTACAGAGATCAACTAGACTTTGAAATATCGCACGCAATTTGGC 1332
Db 2196 GATATGACTATTGTATAGCGGATTTTATCAACACCTAAATATAGCTTAGGGAATATGTA 2255

Qy 1333 ATTTTGAAGATGGAAGGATGGTGTACCACTGCGACGATATAAAGGAAT 1382
Db 2256 TGGTATGACACTAATGAAGATGGTATTCAAGGTGATGATGAAGGAAT 2305

RESULT 9
US-09-386-959-7
; Sequence 7, Application US/093869597
; Patent No. 6703025
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M.
; APPLICANT: FOSTER, Timothy J.
; APPLICANT: HOOK, Magnus
; TITLE OF INVENTION: MULTICOMPONENT VACCINES
; FILE REFERENCE: P06333US1/BAS
; CURRENT APPLICATION NUMBER: US/09/386,959
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: 60/098,439
; EARLIER FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2976
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4)..(2976)
; US-09-386-959-7

Query Match      2.2%; Score 38.8; DB 3; Length 2976;
Best Local Similarity 51.8%; Pred. No. 0.41;
Matches 88; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 1213 AAAAAGGCTAAGCCCATCCATGGAGCTGATGCTCTTTGAAAGCAATTTAACAATAGATGGT 1272
Db 2136 ACAATCTGCTAGACTCAGAAGGCAATTTCTGTATGGGTAACTATTAAACGGACAAGAC 2195

Qy 1273 GATGTGGGTATTTCAGTACAGAGATCAACTAGACTTTGAAATATCGCACGCAATTTGGC 1332
Db 2196 GATATGACTATTGTATAGCGGATTTTATCAACACCTAAATATAGCTTAGGGAATATGTA 2255

Qy 1333 ATTTTGAAGATGGAAGGATGGTGTACCACTGCGACGATATAAAGGAAT 1382
Db 2256 TGGTATGACACTAATGAAGATGGTATTCAAGGTGATGATGAAGGAAT 2305
```

```
; Sequence 7, Application US/093869597
; Patent No. 6703025
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M.
; APPLICANT: FOSTER, Timothy J.
; APPLICANT: HOOK, Magnus
; TITLE OF INVENTION: MULTICOMPONENT VACCINES
; FILE REFERENCE: P06333US1/BAS
; CURRENT APPLICATION NUMBER: US/09/386,959
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: 60/098,439
; EARLIER FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2976
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4)..(2976)
; US-09-386-959-7

Query Match      2.2%; Score 38.8; DB 3; Length 2976;
Best Local Similarity 51.8%; Pred. No. 0.41;
Matches 88; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 1213 AAAAAGGCTAAGCCCATCCATGGAGCTGATGCTCTTTGAAAGCAATTTAACAATAGATGGT 1272
Db 2136 ACAATCTGCTAGACTCAGAAGGCAATTTCTGTATGGGTAACTATTAAACGGACAAGAC 2195

Qy 1273 GATGTGGGTATTTCAGTACAGAGATCAACTAGACTTTGAAATATCGCACGCAATTTGGC 1332
Db 2196 GATATGACTATTGTATAGCGGATTTTATCAACACCTAAATATAGCTTAGGGAATATGTA 2255

Qy 1333 ATTTTGAAGATGGAAGGATGGTGTACCACTGCGACGATATAAAGGAAT 1382
Db 2256 TGGTATGACACTAATGAAGATGGTATTCAAGGTGATGATGAAGGAAT 2305

RESULT 10
US-09-248-796A-2786/c
; Sequence 2786, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 2786
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Candida albicans
; US-09-248-796A-2786

Query Match      2.2%; Score 38.2; DB 3; Length 942;
Best Local Similarity 54.7%; Pred. No. 0.31;
Matches 76; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 751 ATACTAGAGATGATATGGAATTTGCCCTCGATTTTGTGACTTTTGTGAGCTGGAGCT 810
Db 196 ATTCAAATGTTGATATCGTTTGGCCATGATTTTCTGATGTGTTGACTGTGGGGT 137

Qy 811 ACTCTTCTTGACAGACAAAGTCGATTATGGCTATTTCTTTGGAATGACAAATGACAA 870
Db 136 GTTCTCTTGATCAATAAAGAAATTGTATAGGAGATTTATTATTATAAAGCGATGCATCAT 77
```

QY 871 ATGCAGTTTGCCAGATC 889
Db 76 AATCTTTATGAATAAATC 58

RESULT 11

US-09-949-016-16440/C
; Sequence 16440, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16440
; LENGTH: 32172
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(32172)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16440

Query Match 2.1%; Score 36.8; DB 3; Length 32172;
Best Local Similarity 48.1%; Pred. No. 7.8;
Matches 104; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
QY 301 CAGATTAGCTGCAGCAAGAAATAGTTGCTCTTGAAGAACAATGAAGCGTCAGGAC 360
Db 18939 CAGATCATTTTGACTCTTGTCCTAGAAATGTACTTGGCACATAGCAGAACTCAAGCAC 18880
QY 361 CAGGAGTCCGACAAATTAAGGCTCTTTGTCAGATCTTGAAGTAAAGGCATAAAAAG 420
Db 18879 ATTTATTAGTGAATAAATTTGCTGCTTTCTATAGTAATAAACAATTCATAAGAAATG 18820
QY 421 TTGATCGAAATGTACAGATGCCAGTGGCTGCTGTAGTTGTATGGCTTGCAATCGGGCT 480
Db 18819 TTTAACTTAATGTATAACTGTCTTAACAGCTTATGTGTTATGTGGATTAGTAGGCA 18760
QY 481 GATTACCTGGAAGAAGACTATTAATCATCTTAAAA 516
Db 18759 GGTATATTGAATAGGTTAATTTATCAGCTAGAAAA 18724

RESULT 12

US-09-620-312D-194
; Sequence 194, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yunging
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghaast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 194
; LENGTH: 2353
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (153)..(2309)
US-09-620-312D-194

Query Match 2.1%; Score 36.6; DB 3; Length 2353;
Best Local Similarity 49.2%; Pred. No. 1.8;
Matches 96; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 744 TGTATCATACTAGAAAGATGATGAAATTCGCCCTGATTTTGTGAGGC 803
Db 1316 TGCTGTGTTCTGGAAGAGACCTGGACATTCGTGGATTTTTCAGTTTCTTGAGCCA 1375
QY 804 TGGAGCTACTCTTTTGACAGACAAAGTCGATTTATGGCTATTTCTTTGGAATGCAA 863
Db 1376 ATCATCCACCTACTCGAGGAGGATGACAGCCTGTACTGTCATCTCTGCTGGAATGACCA 1435
QY 864 TGACAAATGTCAGTTTGTCCAAGATCCTTATGCTCTTTACCGCTCAGATTTTTCCTGG 923
Db 1436 GGGGTATGAACACACACCGCTGAGGACCCAGCAGCTACTGTACCGTGTGGAGACCATGCGCTGG 1495
QY 924 TCTTGATGGATGCT 938
Db 1496 GCTGGCTGGTGCT 1510

RESULT 13

US-10-012-231A-87
; Sequence 87, Application US/10012231A
; Patent No. 6924355
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; Acids Encoding the Same
; FILE REFERENCE: P2830PIC23
; CURRENT APPLICATION NUMBER: US/10/012,231A
; CURRENT FILING DATE: 2002-06-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 87

QY 924 TCTTGGATGGATGCT 938
Db 1473 GCTGGGCTGGGTGCT 1487

RESULT 15
US-10-006-768A-87
; Sequence 87, Application US/10006768A
; Patent No. 6936697
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C10
; CURRENT APPLICATION NUMBER: US/10/006,768A
; CURRENT FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 477
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 87
; LENGTH: 2725
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-006-768A-87

Query Match 2.1%; Score 36.6; DB 3; Length 2725;
Best Local Similarity 49.2%; Pred. No. 2;
Matches 96; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 744 TGTATCATCTAGAGATGATATGGAATTCGCCCTGATTTTGTGACTTTTGTGAGGC 803
Db 1293 TGCTGTGGTCTTGGAGAGGACCTGGACATTCGTGGATTTTTCAGTTTCTCTGAGCCA 1352

QY 804 TGGAGCTACTCTTCTTGACAGAGACAAGTCGATTATGGCTATTTCTTCTTGAATGACAA 863
Db 1353 ATCCATCCACCTACTCTGGAGAGGATGACAGCCTGTACTGCACTCTGCTGGATGACCA 1412

QY 864 TGGACAAATGCAATGCTTTCACAGATCCTTATGCTCTTTACCGCTCAGATTTTTTCCCGG 923
Db 1413 GGGGTATGAACACACCGCTGAGGACCCAGCACTACTGTACCGTGTGGAGACCATGCTGG 1472

QY 924 TCTTGGATGGATGCT 938
Db 1473 GCTGGGCTGGGTGCT 1487

Search completed: December 14, 2005, 13:08:39
Job time : 343 secs

QY 924 TCTTGGATGGATGCT 938
Db 1473 GCTGGGCTGGGTGCT 1487

RESULT 14
US-10-015-389A-87
; Sequence 87, Application US/10015389A
; Patent No. 6936436
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan I.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C48
; CURRENT APPLICATION NUMBER: US/10/015,389A
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 477
; SEQ ID NO 87
; LENGTH: 2725
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-015-389A-87

Query Match 2.1%; Score 36.6; DB 3; Length 2725;
Best Local Similarity 49.2%; Pred. No. 2;
Matches 96; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 744 TGTATCATCTAGAGATGATATGGAATTCGCCCTGATTTTGTGACTTTTGTGAGGC 803
Db 1293 TGCTGTGGTCTTGGAGAGGACCTGGACATTCGTGGATTTTTCAGTTTCTCTGAGCCA 1352

QY 804 TGGAGCTACTCTTCTTGACAGAGACAAGTCGATTATGGCTATTTCTTCTTGAATGACAA 863
Db 1353 ATCCATCCACCTACTCTGGAGAGGATGACAGCCTGTACTGCACTCTGCTGGATGACCA 1412

QY 864 TGGACAAATGCAATGCTTTCACAGATCCTTATGCTCTTTACCGCTCAGATTTTTTCCCGG 923
Db 1413 GGGGTATGAACACACCGCTGAGGACCCAGCACTACTGTACCGTGTGGAGACCATGCTGG 1472

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 06:14:15 ; Search time 1005 Seconds
(without alignments)
11518.976 Million cell updates/sec

Title: US-10-615-144-3
Perfect score: 1737
Sequence: 1 gaattcgccgcattga.....ataatcgccgcgaattc 1737

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 21.*
1: geneseqn1980s.*
2: geneseqn1990s.*
3: geneseqn2000s.*
4: geneseqn2001as.*
5: geneseqn2001bs.*
6: geneseqn2002as.*
7: geneseqn2002bs.*
8: geneseqn2003as.*
9: geneseqn2003bs.*
10: geneseqn2003cs.*
11: geneseqn2003ds.*
12: geneseqn2004as.*
13: geneseqn2004bs.*
14: geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1737	100.0	1737	2 AAX78002	Aax78002 Tobacco G
2	1255	72.3	1669	2 AAX78001	Aax78001 Potato Gn
3	770.4	44.4	1639	3 AAC39540	Aac39540 Arabidops
4	748	43.1	1641	3 AAC50197	Aac50197 Arabidops
5	714.2	41.1	1836	13 ADX54592	Adx54592 Plant ful
6	661.8	38.1	1882	13 ADO81767	Ado81767 Plant ful
7	661	38.1	1211	13 ADR61718	Adr61718 Cotton cd
8	658.6	37.9	2180	10 ADD68118	Add68118 Glycosylt
9	658.6	37.9	9240	10 ADD68117	Add68117 Glycosylt
10	658	37.9	1319	10 ADD68120	Add68120 Glycosylt
11	656.6	37.8	1119	10 ADD68136	Add68136 Plant Gnt
12	581.2	33.5	1898	13 ADT20185	Adt20185 Plant cDN
13	425.8	24.5	896	14 AEB65577	Aeb65577 Rice geno
14	334	19.2	510	2 AAX78003	Aax78003 A. thalia
15	182	10.5	278	7 ADS70553	Ads70553 Corn seed
16	181.8	10.5	2580	11 ACN44097	Acn44097 Mouse mRN
17	181.8	10.5	31477	11 ACN44096	Acn44096 Mouse gen
18	181.6	10.5	2546	10 ADB58546	Adb58546 Toxicity-
19	181.6	10.5	2546	10 ADB53154	Adb53154 Primary r

20	181.6	10.5	2557	2 AAQ62625	Aaq62625 Rat N-ace
21	165.6	9.5	2485	2 AAQ25594	Aaq25594 Rabbit Gn
22	146.4	8.4	1335	4 AA169837	Aai69837 Human bet
23	146.4	8.4	2518	13 ADT07465	Adt07465 Human col
24	146.4	8.4	2602	6 ABK84032	Abk84032 Human cDN
25	146.4	8.4	2614	13 ADT07459	Adt07459 Human col
26	146.4	8.4	2631	13 ADT07458	Adt07458 Human col
27	146.4	8.4	2670	12 ADJ57449	Adj57449 Human Glc
28	146.4	8.4	2670	14 ADZ69846	Adz69846 Human Gnt
29	146.4	8.4	2680	13 ADT07460	Adt07460 Human col
30	146.4	8.4	2688	10 ADI62841	Adi62841 Human apo
31	146.4	8.4	2688	10 ADI62862	Adi62862 Human apo
32	146.4	8.4	2753	11 ACN44099	Acn44099 Human mRN
33	146.4	8.4	2842	13 ADT07454	Adt07454 Human col
34	146.4	8.4	2900	13 ADT07464	Adt07464 Human col
35	146.4	8.4	2936	13 ADT07457	Adt07457 Human col
36	146.4	8.4	2937	13 ADT07430	Adt07430 Human Gnt
37	146.4	8.4	2937	14 ADY14405	Ady14405 DNA encod
38	146.4	8.4	2937	14 ADZ49000	Adz49000 Insulin s
39	146.4	8.4	3067	13 ADT07463	Adt07463 Human col
40	146.4	8.4	3106	13 ADT07462	Adt07462 Human col
41	146.4	8.4	3135	13 ADT07455	Adt07455 Human col
42	146.4	8.4	3183	13 ADT07456	Adt07456 Human col
43	146.4	8.4	3230	2 AAQ25595	Aaq25595 Human Gnt
44	146.4	8.4	4677	12 ADJ57451	Adj57451 Vector pp
45	146.4	8.4	4677	14 ADZ69848	Adz69848 Plasmid p

ALIGNMENTS

RESULT 1
AAX78002
ID AAX78002 standard; cDNA; 1737 BP.

XX AAX78002;

DT 19-AUG-1999 (first entry)

DE Tobacco GntI cDNA.

XX GntI; beta-1,2-N-acetylglucosaminyltransferase; glycoprotein; plant;
KW deficient; defective; detection; transgenic plant; sugar residue;
KW medicine; antigenicity; deglycosylation; tobacco; ds.
XX Nicotiana tabacum.

XX Key Location/Qualifiers
FH 127..1467
CDS /*tag= a

FT /product= "GntI"

XX DE19754622-A1.

XX 10-JUN-1999.

XX 09-DEC-1997; 97DE-01054622.

XX 09-DEC-1997; 97DE-01054622.

XX (VSCH/) VON SCHAEWEN A.

XX Von Schawen A;

XX WPI; 1999-338905/29.

XX P-PSDB; AAY08889.

XX Nucleic acid encoding plant N-acetylglucosaminyl transferase I useful for
generating plants producing glycosylated proteins of low antigenicity.

XX Claim 7; Page 22-24; 37pp; German.

XX This invention describes novel plant N-acetylglucosaminyl transferase I

= priority doc
in invent
appl.

(GnTI) proteins and their encoding nucleic acids. The nucleic acid sequences of the invention may be used for recombinant production of the encoded proteins, which are then used to raise antibodies (Ab) for identifying plants with deficient or defective production of the protein. They may also be used to detect such plants by hybridization and to isolate related sequences from other plants or to generate antisense or sense constructs for reducing/deleting GnTI protein activity in plants. These transgenic plants may be used to produce glycoproteins with minimal, uniform and defined sugar residues. Such glycoproteins are useful in medicine and research, e.g. human glucocerebrosidase for treating Gaucher's disease. Plants which are defective or deficient in production of the GnTI protein can be made to produce glycoproteins with minimal, uniform and defined sugar residues, of low antigenicity. Use of these plants eliminates the need for the difficult isolation and deglycosylation of native proteins or preparation in defective animal cells. This sequence represents tobacco (Nicotiana tabacum) GnTI encoding cDNA

Query Match 100.0%; Score 1737; DB 2; Length 1737;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1737; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGGCGCCGATTTGACTGCTTAACCTCAAGCAGCAAGTAATCCAGCGATGA 60
DB 1 GAATTCGGCGCCGATTTGACTGCTTAACCTCAAGCAGCAAGTAATCCAGCGATGA 60
QY 61 AACACTCAATCACTGAACACTGAGAGACTATTCGGCTTCTCTTAAGCGCTTCAATCGAATT 120
DB 61 AACACTCAATCACTGAACACTGAGAGACTATTCGGCTTCTCTTAAGCGCTTCAATCGAATT 120
QY 121 CGCAGATGAGAGGAAACAAGTTTGTGTGATTTCCGGTACCTCTCATCTTGGCTGCT 180
DB 121 CGCAGATGAGAGGAAACAAGTTTGTGTGATTTCCGGTACCTCTCATCTTGGCTGCT 180
QY 181 GTCGCTTCATCTACACAGATCGGCTTTTCGGACACAGTCAGATATGACATCGC 240
DB 181 GTCGCTTCATCTACACAGATCGGCTTTTCGGACACAGTCAGATATGACATCGC 240
QY 241 CTTGCTGCTCAATTTGAAGCAGAAATCATTTGTAAGCCAGACCCAGATTGCTTATTGAC 300
DB 241 CTTGCTGCTCAATTTGAAGCAGAAATCATTTGTAAGCCAGACCCAGATTGCTTATTGAC 300
QY 301 CAGATTAGCTGCGACGAGGAAGTATGTTCTCTTGAAGAACAAATGAAGCGTCAGGAC 360
DB 301 CAGATTAGCTGCGACGAGGAAGTATGTTCTCTTGAAGAACAAATGAAGCGTCAGGAC 360
QY 361 CAGAGTGCCGACAAATTAAGGGCTCTTGTTCAGGATCTTGAAGTAAGGGCATAAAGAAG 420
DB 361 CAGAGTGCCGACAAATTAAGGGCTCTTGTTCAGGATCTTGAAGTAAGGGCATAAAGAAG 420
QY 421 TTGATCGGAATGTACAGATCCAGTGGCTGCTGTAGTTGTTATGGCTTGCATTCGGGCT 480
DB 421 TTGATCGGAATGTACAGATCCAGTGGCTGCTGTAGTTGTTATGGCTTGCATTCGGGCT 480
QY 481 GATTAACCTGGAAGAGACTATTAATTCATCTTAAATACCAATPATCTGTTCGGTCAAAA 540
DB 481 GATTAACCTGGAAGAGACTATTAATTCATCTTAAATACCAATPATCTGTTCGGTCAAAA 540
QY 541 TATCCTCTTTTCATATCCAGGATGGATCATCTCTGATGTCAGGAGCTTGTTCGAGC 600
DB 541 TATCCTCTTTTCATATCCAGGATGGATCATCTCTGATGTCAGGAGCTTGTTCGAGC 600
QY 601 TATGATCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGCACTATGAAAGACCA 660
DB 601 TATGATCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGCACTATGAAAGACCA 660
QY 661 GGGAGCTGATTCATATCAAAAATTGACGCTCATTTACAAAGTGGCGATTGGATTCAGCTG 720
DB 661 GGGAGCTGATTCATATCAAAAATTGACGCTCATTTACAAAGTGGCGATTGGATTCAGCTG 720
QY 721 TTTTACAAGCATATTTTAGCCGTGTTATCATACTAGNAGATGATATGGAATTTGCCCT 780

DB 721 TTTTACAAGCATATTTTAGCCGTGTTATCATACTAGAAGATGATATGGAATTTGCCCT 780
QY 781 GATTTTTTTGACTTTTTTTGAGGCTGAGCTACTCTTTTGCACAGACAAAGTCGATTATG 840
DB 781 GATTTTTTTGACTTTTTTTGAGGCTGAGCTACTCTTTTGCACAGACAAAGTCGATTATG 840
QY 841 GCTATTTCTTCTTGGATGACAAATGCAAAATGCAAGTTTGTCCAAAGATCCTTATGCTCT 900
DB 841 GCTATTTCTTCTTGGATGACAAATGCAAAATGCAAGTTTGTCCAAAGATCCTTATGCTCT 900
QY 901 TACCGCTCAGATTTTTTCCCGGTCTTGTGATGATGCTTTTCAAAATCTACTTCGGACGAA 960
DB 901 TACCGCTCAGATTTTTTCCCGGTCTTGTGATGATGCTTTTCAAAATCTACTTCGGACGAA 960
QY 961 TTATCTCCAAAGTGGCCAAAGGCTTACTCGGACGACTGAGCTCAAGAGAAATCAC 1020
DB 961 TTATCTCCAAAGTGGCCAAAGGCTTACTCGGACGACTGAGCTCAAGAGAAATCAC 1020
QY 1021 AGAGTTCGACAAATTTATTCGCCAGAAAGTTTGCAGAAACATATAATTTTGGTGAAGCATG 1080
DB 1021 AGAGTTCGACAAATTTATTCGCCAGAAAGTTTGCAGAAACATATAATTTTGGTGAAGCATG 1080
QY 1081 TCTAGTTTTCGGGCGAGTTTTTCAAGCAGATATCTTGAGCCAAATTAACCTAAATGATGTCAG 1140
DB 1081 TCTAGTTTTCGGGCGAGTTTTTCAAGCAGATATCTTGAGCCAAATTAACCTAAATGATGTCAG 1140
QY 1141 GTTGATTGGAAGTCAATGGACCTTAGTTACTCTTTTGGAGGACAAATTTACGTGAACACTTT 1200
DB 1141 GTTGATTGGAAGTCAATGGACCTTAGTTACTCTTTTGGAGGACAAATTTACGTGAACACTTT 1200
QY 1201 GGTGACTTGGTTAAAGGCTAAGCCCATCCATGGAGCTGATGCTCTCTTGAAGCAATTT 1260
DB 1201 GGTGACTTGGTTAAAGGCTAAGCCCATCCATGGAGCTGATGCTCTCTTGAAGCAATTT 1260
QY 1261 AACATAGATGTTGATGCTGCTATTCAGTACAGAGATCAACTGACCTTTGAAATATCGCA 1320
DB 1261 AACATAGATGTTGATGCTGCTATTCAGTACAGAGATCAACTGACCTTTGAAATATCGCA 1320
QY 1321 CGGCAATTTGGCATTTTTTGAAGAAATGGAAGGATGGTGATACCAGTGCAGCATATAAAGGA 1380
DB 1321 CGGCAATTTGGCATTTTTTGAAGAAATGGAAGGATGGTGATACCAGTGCAGCATATAAAGGA 1380
QY 1381 ATAGTAGTTTTCGGTACAAAGCTCCAGCGTGATATTCCTTGTGGCCATGATTCGCTT 1440
DB 1381 ATAGTAGTTTTCGGTACCAAGCTCCAGCGTGATATTCCTTGTGGCCATGATTCGCTT 1440
QY 1441 CAACAACTCGGAATTCGAAGATCTTAACAAAGATATGATTGCAAGGAGCCCGGCAAAATTT 1500
DB 1441 CAACAACTCGGAATTCGAAGATCTTAACAAAGATATGATTGCAAGGAGCCCGGCAAAATTT 1500
QY 1501 TTTGACTTATTTGGGTAGGATGCATCAGCTGCACCTAAACCATGATTTTACCAAGTTACAT 1560
DB 1501 TTTGACTTATTTGGGTAGGATGCATCAGCTGCACCTAAACCATGATTTTACCAAGTTACAT 1560
QY 1561 ACAACGTTTAAATGTTATACGGAGGAGCTCAGTCTTCTAGTGTGTAAGGGATATTCGGCTT 1620
DB 1561 ACAACGTTTAAATGTTATACGGAGGAGCTCAGTCTTCTAGTGTGTAAGGGATATTCGGCTT 1620
QY 1621 CTTAGTATTTGGATGATCATCAACCAACTATTTATTTAAAGTGTTCAGAACATAAAGAG 1680
DB 1621 CTTAGTATTTGGATGATCATCAACCAACTATTTATTTAAAGTGTTCAGAACATAAAGAG 1680
QY 1681 GAAATGAGCCCTGTAAAGACTATACATGGGACCATCATTAATCGCGGCCGCGAATTC 1737
DB 1681 GAAATGAGCCCTGTAAAGACTATACATGGGACCATCATTAATCGCGGCCGCGAATTC 1737

RESULT 2
AA78001
ID AAX78001 standard; cDNA; 1669 BP.
XX
AC AAX78001;

XX	19-AUG-1999	(first entry)	QY	306	TAGCTGCAGCAAGGAAGAAATAGTGTCTTGTGAAGAA	CAAAATGAAGCGTCAGGACCAAGGA	365
DT			DB	232	TAGCCACGACGAGGAGAGTAGTCTTTGAAGAACAAATGAAGCATCAGGACCAAGGA	291	
DE	Potato GntI cDNA.		QY	366	GTGCGCAATTAAGGGCTCTTGTGAGGATCTTGAAGTAAGGCAATAAAAAGTTGAT	425	
KW	GntI; beta-1,2-N-acetylglucosaminyltransferase; glycoprotein; plant;		DB	292	GTGCGGCAATTAAGGGCTCTTGTGAGGATCTTGAAGTAAGGCAATAAAAAGTTAAT	351	
KW	deficient; detection; transgenic plant; sugar residue;		QY	426	CGGAAATGTACAGATGCCAGTGGCTGTGTAGTTGTTATGCTTTCGAATCGGGTGATTA	485	
XX	medicine; antigenicity; deglycosylation; potato; ds.		DB	352	CGGAGATGTGCAGATGCCAGTGGCAGCTGTAGTTGTTATGCTTTCGACGTCTGACTACTA	411	
OS	Solanum tuberosum.		QY	486	CCTGGAAAGAGCTATTAATCCATCTTAAATACCAATATCTCTTGGGTCAAAATATCC	545	
FH	Key	Location/Qualifiers	DB	412	CCTGGAGGAGCTATTAATCCATCTTAAATACCAATATCTCTTGGGTCAAAATATCC	471	
FT	53..1394	/*tag= a	QY	546	TCCTTTCATATCCAGGATGCATCATCTCTGATGTCTGAGGAGCTGCTTGTGAGCTATGA	605	
FT	/product= "GntI"		DB	472	TCCTTTCATATCCAGGATGCATCATCTCTGATGTCTGAGGAGCTGCTTGTGAGCTATGA	531	
XX	DE19754622-A1.		QY	606	TCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGTGCATATCTGAAAACCAAGGGA	665	
PN	10-JUN-1999.		DB	532	TCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGTGCATATCTGAAAACCAAGGGA	591	
XX	09-DEC-1997; 97DE-01054622.		QY	666	GCTGATTGCATATACAAAATTGACGTCATTAACAAGTGGGCATTTGATAGCTGTTTTA	725	
XX	09-DEC-1997; 97DE-01054622.		DB	592	ACTGTTGCATATACAAAATTGACGTCATTAACAAGTGGGCATTTGATAGCTGTTTTCA	651	
XX	(VSCH/) VON SCHAEWEN A.		QY	726	CAAGCATTAATTTAGCCGTGTATCATCTACTAGAGATGATATGAAAATTTGCCCTGATTT	785	
XX	Von Schaewen A;		DB	652	CAAGCATTAATTTAGCCGTGTATCATCTACTAGAGATGATATGAAAATTTGCCCTGATTT	711	
XX	WPI; 1999-338905/29.		QY	786	TTTTGACTTTTTGAGGCTGAGCTACTCTTCTTTCGACAGACCAAGTCGATATGCGCTAT	845	
DR	P-PSDB; AAY08888.		DB	712	TTTTGACTTTTTGAGGCTGAGCTACTCTTCTTTCGACAGACCAAGTCGATATGCGCTAT	771	
PT	Nucleic acid encoding plant N-acetylglucosaminyl transferase I useful for		QY	846	TTCTTCTTGGAAATGACAAATGACAAATGCGATTTGTTCGAAATCTTATGCTCTTTACCG	905	
PT	generating plants producing glycosylated proteins of low antigenicity.		DB	772	TTCTTCTTGGAAATGACAAATGACAAATGCGATTTGTTCGAAATCTTATGCTCTTTACCG	831	
XX	Claim 6; Page 16-18; 37pp; German.		QY	906	CTCAGATTTTTTCCCGTCTTGGATGATCTCTTTCGACAGACCAAGTCGATATGCGCTAT	965	
CC	This invention describes novel plant N-acetylglucosaminyl transferase I		DB	832	CTCAGATTTTTTCCCGTCTTGGATGATCTCTTTCGACAGACCAAGTCGATATGCGCTAT	891	
CC	(GntI) proteins and their encoding nucleic acids. The nucleic acid		QY	966	TCCAAAGTGGCCAAAGGCTTACTGGGACGCTGGCTAAGACTTCAAAAGAGATTCACAGAGG	1025	
CC	sequences of the invention may be used for recombinant production of the		DB	892	TCCAAAGTGGCCAAAGGCTTACTGGGACGCTGGCTAAGACTTCAAAAGAGATTCACAGAGG	951	
CC	encoded proteins, which are then used to raise antibodies (Ab) for		QY	1026	TCGACAAATTTATTCGCCCCAGAAAGTTTGCGAAACGTCACAAATTTTGGTGAGCATGGTTCTAG	1085	
CC	identifying plants with deficient or defective production of the protein.		DB	952	TCGACAAATTTATTCGCCCCAGAAAGTTTGCGAAACGTCACAAATTTTGGTGAGCATGGTTCTAG	1011	
CC	They may also be used to detect such plants by hybridization and to		QY	1086	TTTGGGGCAGTTTTTCAAGCAGATCTTTCGACCAATTAACCTAAATGATGCCAGTTGA	1145	
CC	isolate related sequences from other plants or to generate antisense or		DB	1012	TTTGGGGCAGTTTTTCAAGCAGATCTTTCGACCAATTAACCTAAATGATGCCAGTTGA	1071	
CC	sense constructs for reducing/deleting GntI protein activity in plants.		QY	1146	TTTGGGGCAGTTTTTCAAGCAGATCTTTCGACCAATTAACCTAAATGATGCCAGTTGA	1205	
CC	These transgenic plants may be used to produce glycoproteins with		DB	1072	TTTGGGGCAGTTTTTCAAGCAGATCTTTCGACCAATTAACCTAAATGATGCCAGTTGA	1131	
CC	minimal, uniform and defined sugar residues. Such glycoproteins are		QY	1206	CTTGTGTTAAAAAGGCTTAAGCCCATCCATGGAGCTGATCTGTCTTGAAGCAATTTAAACAT	1265	
CC	useful in medicine and research, e.g. human glucocerebrosidase for		DB	1132	CTTGTGTTAAAAAGGCTTAAGCCCATCCACGAGCTGATCTGTCTTGAAGCAATTTAAACAT	1191	
CC	treating Gaucher's disease. Plants which are defective or deficient in		QY	1266	AGATGGTGATGTGGCTTATTCAGTACAGAGATCACTAGACTTTGAAAATATCGCAGCGCA	1325	
CC	production of the GntI protein can be made to produce glycoproteins with		DB	1192	AGATGGTGATGTGGCTTATTCAGTACAGAGATCACTAGACTTTGAAAATATCGCAGCGCA	1251	
CC	minimal, uniform and defined sugar residues, of low antigenicity. Use of		QY	1326	ATTTGGCATTTTTTGAAGAAATGGAAGATGGTGATCCAGTGCAGCATATAAAGGAATAGT	1385	
CC	these plants eliminates the need for the difficult isolation and		DB	1252	GTTCGGCATTTTTTGAAGAAATGGAAGATGGTGATCCAGTGCAGCATATAAAGGAATAGT	1311	
CC	deglycosylation of native proteins or preparation in defective animal		QY	1386	AGTTTTTCGGTACCAAAAGTCCAGACGCTGTATTTCTTGTGTGCGCATGATTCGCTTCAACA	1445	
CC	cells. This sequence represents potato (Solanum tuberosum) GntI encoding						
CC	cDNA						
XX	Sequence 1669 BP; 489 A; 312 C; 387 G; 481 T; 0 U; 0 Other;		SQ				
XX	Query Match	72.3%; Score 1255; DB 2; Length 1669;					
XX	Best Local Similarity	89.1%; Pred. NO. 0;					
XX	Matches 1412; Conservative	0; Mismatches 140; Indels 33; Gaps 4;					
QY	126	GATGAGAGGAAACAGTTTTTGTGTCATTTCCGGTACCTCCATCTTGGCTGCTGCTCGC	185				
DB	52	GATGAGAGGAAACAGTTTTTGTGTCATTTTACGGTACCTTCGCTGCTGCTCTCGC	111				
QY	186	CTTCATCTACACACAGATCGGCTTTTTCGACACAGTCAGATATGAGATCGCCTTGC	245				
DB	112	CTTCATCTACACAGATCGGCTTTTTCGACACAGTCAGATATGAGATCGCCTTGC	171				
QY	246	TGCTGCAATTTGAAGCAGAAAATCATTGTACAAGCCAGACAGATTCCTTATTGACAGAT	305				
DB	172	TGCTGCAATTTGAAGCAGAAAATCATTGTACAAGCCAGACAGATTCCTTATTGACAGAT	231				

priority doc
in instant appl.

PR	27-JUL-1999;	99US-01459139P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161922P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	08-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	11-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157753P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158369P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159331P.			
PR	14-OCT-1999;	99US-0159637P.			
PR	14-OCT-1999;	99US-0159638P.			
PR	18-OCT-1999;	99US-0159584P.			
PR	21-OCT-1999;	99US-0160741P.			
PR	21-OCT-1999;	99US-0160767P.			
PR	21-OCT-1999;	99US-0160768P.			
PR	21-OCT-1999;	99US-0160770P.			
PR	21-OCT-1999;	99US-0160814P.			
PR	21-OCT-1999;	99US-0160815P.			
PR	22-OCT-1999;	99US-0160980P.			
PR	22-OCT-1999;	99US-0160981P.			
PR	22-OCT-1999;	99US-0160989P.			
PR	25-OCT-1999;	99US-0161404P.			
PR	25-OCT-1999;	99US-0161405P.			
PR	25-OCT-1999;	99US-0161406P.			
PR	27-JUL-1999;	99US-01459139P.	PR	2	

Db	938	CGACCGGAAGTCGTAGAACATACAAATTTTGGTGAACATGGTCTAGTTTGGGACAGTTT	997	PR	06-MAY-1999;	99US-0132486P.
Qy	1099	TTCAAGCAGTATCTTGAGCCAAATTAACATAATGATGCCAGGTGATGAGTGAAGTCAATG	1158	PR	06-MAY-1999;	99US-0132487P.
Db	998	TTCAAGCAGTATCTTGAGCCAAATTAACATAATGATGCCAGGTGATGAGTGAAGTCAATG	1057	PR	07-MAY-1999;	99US-0132863P.
Qy	1159	GACCTTAGTTACCTTTTGGAGGACAAATTAACGTGAACACATTTTGGTGACCTTGGTTTAAAG	1218	PR	11-MAY-1999;	99US-0134256P.
Db	1058	GACCTGGATACCTTGACAGAGGGAACATAACCAAGTACTTTCTGGCTTAGTGAGACAA	1117	PR	14-MAY-1999;	99US-0134218P.
Qy	1219	GCTAAGCCCATCCATGGAGCTGATGCTGTCTTTGAAAGCATTTAAACATAGATGGTGAATG	1278	PR	14-MAY-1999;	99US-0134219P.
Db	1118	GCACGACCAATTAAGGTTCTGACCTTGCTTTAAAGGCTCAAAACATAAAGGATGATTT	1177	PR	14-MAY-1999;	99US-0134221P.
Qy	1279	CGTATTCAGTACAGAGATCAACTAGACTTTGAAAATATCGCACGGCAATTTGGCATTTTT	1338	PR	18-MAY-1999;	99US-0134370P.
Db	1178	CGTATCCGGTATAAAGACCAAGTAGAGTTTGAACGCATTTGCAGGGGAATTTGGTATATTT	1237	PR	18-MAY-1999;	99US-0134376P.
Qy	1339	GAAGAAATGGAAGGATGGTGACCAAGTGCAGCATATAAAGGAATAGTATTTCCGGTAC	1398	PR	20-MAY-1999;	99US-0135124P.
Db	1238	GAAGAAATGGAAGGATGGTGACCAAGTGCAGCATATAAAGGAATAGTATTTCCGGTAC	1297	PR	21-MAY-1999;	99US-0135353P.
Qy	1399	CAACGCTCCAGACGTGATTTCTTTGGCCCATGATTCGCTTCAACCACTCGGAATTTGAA	1458	PR	24-MAY-1999;	99US-0135629P.
Db	1298	CAGACAAACACGCTGATTTCTTGTTGGGCCAGATTCGTAAATGCAGCTTGGAAATCGA	1357	PR	25-MAY-1999;	99US-0136021P.
Qy	1459	GATACTTAAACAAAGATGATGTCAGGA	1486	PR	27-MAY-1999;	99US-0136392P.
Db	1358	AATTCCTGATCAAAACATATGAAAGGA	1385	PR	28-MAY-1999;	99US-0136782P.
RESULT 4						99US-0137222P.
ID	AAC50197 standard; DNA; 1641 BP.					99US-0137528P.
XX	AAC50197;					99US-0137502P.
AC	AAC50197;					99US-0138540P.
XX	18-OCT-2000 (first entry)					99US-0138847P.
DT	Arabidopsis thaliana DNA fragment SEQ ID NO: 63936.					99US-0139119P.
DE	Hybridisation assay; Genetic mapping; gene expression control;					99US-0139452P.
XX	protein identification; signal transduction pathway; metabolic pathway;					99US-0139453P.
KW	promoter; termination sequence; ss.					99US-0139457P.
XX	Arabidopsis thaliana.					99US-0139458P.
OS	EP1033405-A2.					99US-0139459P.
XX	06-SEP-2000.					99US-0139460P.
XX	25-FEB-2000; 2000EP-00301439.					99US-0139461P.
PF	25-FEB-1999;					99US-0139462P.
XX	99US-0121825P.					99US-0139463P.
PR	05-MAR-1999;					99US-0139463P.
PR	09-MAR-1999;					99US-0139750P.
PR	23-MAR-1999;					99US-0139763P.
PR	25-MAR-1999;					99US-0139817P.
PR	29-MAR-1999;					99US-0139899P.
PR	01-APR-1999;					99US-0140353P.
PR	06-APR-1999;					99US-0140354P.
PR	08-APR-1999;					99US-0140695P.
PR	16-APR-1999;					99US-0140823P.
PR	21-APR-1999;					99US-0140991P.
PR	23-APR-1999;					99US-0141287P.
PR	30-APR-1999;					99US-0141842P.
PR	04-MAY-1999;					99US-0142055P.
PR	05-MAY-1999;					99US-0142390P.

PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-01452224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	26-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145913P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161922P.
PR	03-AUG-1999;	99US-0147033P.	PR	28-OCT-1999;	99US-0161933P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0161933P.
PR	04-AUG-1999;	99US-0147302P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.	Query Match 43.1%; Score 748; DB 3; Length 1641;		
PR	05-AUG-1999;	99US-0147260P.	Best Local Similarity 73.1%; Pred. No. 6.2e-217;		
PR	06-AUG-1999;	99US-0147303P.	Matches 988; Conservative 0; Mismatches 360; Indels 4; Gaps 2;		
PR	06-AUG-1999;	99US-0147418P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147933P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	28-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157753P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158369P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159331P.			
PR	14-OCT-1999;	99US-0159637P.			
PR	14-OCT-1999;	99US-0159638P.			
PR	18-OCT-1999;	99US-0159584P.			
PR	21-OCT-1999;	99US-0160741P.			
PR	21-OCT-1999;	99US-0160767P.			
PR	21-OCT-1999;	99US-0160768P.			
PR	21-OCT-1999;	99US-0160770P.			
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-01452224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	26-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145913P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161922P.
PR	03-AUG-1999;	99US-0147033P.	PR	28-OCT-1999;	99US-0161933P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0161933P.
PR	04-AUG-1999;	99US-0147302P.	Query Match 43.1%; Score 748; DB 3; Length 1641;		
PR	05-AUG-1999;	99US-0147192P.	Best Local Similarity 73.1%; Pred. No. 6.2e-217;		
PR	05-AUG-1999;	99US-0147260P.	Matches 988; Conservative 0; Mismatches 360; Indels 4; Gaps 2;		
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147418P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147933P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	28-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157753P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158369P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159331P.			
PR	14-OCT-1999;	99US-0159637P.			
PR	14-OCT-1999;	99US-0159638P.			
PR	18-OCT-1999;	99US-0159584P.			
PR	21-OCT-1999;	99US-0160741P.			
PR	21-OCT-1999;	99US-0160767P.			
PR	21-OCT-1999;	99US-0160768P.			
PR	21-OCT-1999;	99US-0160770P.			

Db 816 TTTCTGGCTTGGTGGATGCTCAAGAGATCGATCTGGGATGATGTTATACCAAAATGG 875
Qy 976 CCAAAGG-CTTACTGGGACGACTGGCTAAGACTCAAAAGAGAATCACAGAGTGCACAATT 1034
Db 876 CCAAAGGCTTACTGGGATGATTGGCTGAGACTAAAGGAAACCAATAAGCGCCCAATT 935
Qy 1035 TATTGCCCCAGAAAGTTGCAGAACATATAATTTTGGTGGAGCATGGTCTAGTTTGGGGCA 1094
Db 936 CATTCGACCGGAAGTCTGTAGAACATACAAATTTTGGTGAACATGGGTCTAGTTTGGGACA 995
Qy 1095 GTTTTTCAGGACGATCTGTAGCCAAATTAATACTAAATGATCTCCAGGTGATTGCGAAGTC 1154
Db 996 GTTTTTCAGTCACTGTGGAACCTATAAAGCTAAACGATGACGGTGTACTGGAAAGC 1055
Qy 1155 AATGGACCTTAGTTACCTTTTGGAGGACAAATACGTGAAACACTTTTGGTGACTTGGTTAA 1214
Db 1056 AAGGACCTGGGATACCTGCAGAGAGGAAACTATACCAAGTACTTTTCTGGCTTAGTGAG 1115
Qy 1215 AAGGCTAAGCCCATCCATGAGCTGATGCTGCTTGAAGCATTTAACAATAGATGTGA 1274
Db 1116 ACAAGCAGCAACCAATCAAGGTTCTGACCTTTGTCTTAAAGGCTCAAAACATAAAGGATGA 1175
Qy 1275 TGTGCGTATTCAGTACAGAGATCAACTAGACTTTTGAATAATTCGACGGCAATTTGGCAT 1334
Db 1176 TGTTCGTATCCGTATATAAGACCAAGTAGATTTGAACGCATTCAGGGNAATTTGGTAT 1235
Qy 1335 TTTTGAAGAATGGAAGGATGTTGACCAAGTGCAGCATATAAAGAAATAGTAGTTTTCG 1394
Db 1236 ATTTGAAGAATGGAAGGATGTTGCTCGAACAGCATATAAAGGATAGTGTGTTTCG 1295
Qy 1395 GTACCAACGTCAGACGTGATTCCTTGTGTCGATGATTCGGTTCACAACTCGGAAT 1454
Db 1296 AATCCAGACAAACAGACGTGATTCCTGTTGGTGGCCAGATTCGTGAATGCAGCTTGGAA 1355
Qy 1455 TGAAGATCTTAACAAAGATATGATTCGAGGA 1486
Db 1356 TCGAAATTCCTGTGTCGAAACATATGAAGA 1387

RESULT 5
ID ADX54592 standard; cDNA, 1836 BP.
AC ADX54592;
XX
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polynucleotide seqid 29332.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomanan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
XX protein content; gene; ss.
OS Unidentified.
XX
XX US2004034888-A1.
XX
PD 19-FEB-2004.
XX
XX 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
XX (LIUJ/) LIU J.
XX (ZHOU/) ZHOU Y.
XX (KOVA/) KOVALIC D K.

PA (SCRE/) SCREEN S E.
PA (TAB/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
DR
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 29332; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.sequdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomanan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 1836 BP; 546 A; 326 C; 406 G; 558 T; 0 U; 0 Other;
Query Match 41.1%; Score 714.2; DB 13; Length 1836;
Best Local Similarity 71.8%; Pred. No. 1.4e-206;
Matches 1037; Conservative 0; Mismatches 288; Indels 120; Gaps 3;
Qy 139 AAGTTTGTGTGATTTTCGGGTACCTCCTCATCTTGGCTGCTGCGCTTCATCTACACA 198
Db 43 AAAGTTTCTGTGATTTTCGGTTTCTCTCTGATTCGAGTCGGCTTCATCTATATC 102
Qy 199 CAGATCGGCTTTTTCGGACACAGTCAGATATGAGATGCGCTTGTGTCGCAATTGAA 258
Db 103 CAGATGAGGCTTTTCGCAACACAAATCAGAAATATGCTGCGCTGCGTGCAGCTATCGAA 162
Qy 259 GCAGAAATCATTTGACAGCCAGACAGATTCCTTATTGACAGATTCAGCTGCAGCAA 318
Db 163 GCTGAAACCAATTTGACAAAGTCAACCGCATCGCTGATGATCAGATTAGCTTGCAACA 222
Qy 319 GGAAGAAATAGTTCTCTTGAAGAAACAAAT----- 347
Db 223 GGACGAATTTGCGCCCTGGAAGATTAATTTGGTGGTCTTCTTTCAGAGAGC 282
Qy 348 -GAAGCGTCAGGACAGGAGTCCGCAAAATTAAGGCTCTTGTTCAGGATCTTGAAGTA 406
Db 283 AGAAACGTCGAGACCAAGAATGTGGACAAACAAAGTCTCTTTGACAAAGATCTTTGAAAGAA 342
Qy 407 AGGCGATAAAAGTTGATCGGAAATGTACAGATGCGAGTGGCTGTGTAGTTGTTATGG 466
Db 343 AAGACCTGCAGAGGCTGATTAAGTGAAGTTCGGTGGCTGCTTGTGTGATCATGG 402
Qy 467 CTTTGAATCGGGCTGATTAACCTGGAAAAAGACTATAAATCCATCTTAAAAATACCAATAT 526
Db 403 CATGTAATCGTCTGATTAACCTGGAGAGGACTATAATCTGTATTGAATACCAAGGC 462
Qy 527 CTGTTGGCTCAAAATATCTCTTTTCATATCCAGAGATGATCATCTCTGTATGTACGGA 586
Db 463 CCATTTCTTCAGATATCTCTTTATTTGTATCTCAGGATGGATCAAAATCCAAATGTTAAAA 522
Qy 587 AGCTTGTCTTTGAGCTATGATCAGCTGACGATATATGACAGCACTTGGATTTTGAACCTGTGC 646

[illegible]

Db	1300	ATACGACGACCAAGAAGTTTCGAGCAGATAGCTCTGTCATATTTTGGAAATATTCGAAGAATG	1355
Qy	1347	GAAGGATGGTGTACCCAGCTGCAGCATATAAAGAAATAGTAGTTTTCGGTACCA---AAC	1403
Db	1360	GAAGGATGGCATCCCAAGACGGCTTACAAGAAGGTGGTGCTCTTCGGTACAACAGTAG	1419
Qy	1404	GTCACGACGTGTATTCCTTGTTCGGCCATGANTCGCTTCACAACTCGGAATCGAAGATAC	1463
Db	1420	TCAAAGCGCAATATTTCTCGTCAGCCACAGATTCTCTTCGTGTCAGCTGGGGGTAGCATGT	1479
Qy	1464	TTACAAGATATGATTGTCAGGAGCCCGGCAAAATTTTGTG	1504
Db	1480	TGAAGAGGTAGATTTTACCCTTTGGAATGCGCAATATCATG	1520
RESULT 7			
AD	61718	ID ADR61718 standard; cDNA; 1211 BP.	
XX	AC	ADR61718;	
XX	XX	02-DEC-2004 (first entry)	
DE	XX	Cotton cDNA sequence, SEQ ID 2499.	
XX	XX	Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;	
KW	KW	drought tolerance; plant disease resistance; galactomannan; lignin;	
KW	KW	plant growth regulator; heat tolerance; herbicide tolerance;	
KW	KW	homologous recombination; extreme osmotic condition tolerance;	
KW	KW	pathogen resistance; pest resistance; yield; photosynthesis; seed oil;	
XX	XX	stress resistance.	
OS	OS	Gossypium hirsutum.	
XX	XX	US2004181830-A1.	
XX	XX	16-SEP-2004.	
XX	XX	29-JAN-2004; 2004US-00767795.	
XX	XX	07-MAY-2001; 2001US-00849529.	
PR	PR	12-DEC-2001; 2001US-00021323.	
XX	PA	(KOVA/) KOVALIC D K.	
PA	PA	(ZHOU/) ZHOU Y.	
PA	PA	(CAOY/) CAO Y.	
XX	XX	Kovalic DK, Zhou Y, Cao Y;	
FI	FI	WPI; 2004-667718/65.	
DR	DR	New recombinant nucleic acid molecules and polypeptides from Gossypium	
XX	XX	hirsutum, useful for producing plants with improved biological	
PT	PT	characteristics (e.g. improved plant cold or drought tolerance).	
XX	XX	Claim 1; SEQ ID NO 2499; 14pp; English.	
XX	XX	The invention relates to a recombinant polynucleotide comprising any of	
CC	CC	the 58798 Cotton plant cDNA sequences mentioned in the specification.	
CC	CC	Also a recombinant polypeptide comprising any of the 58798 amino acid	
CC	CC	sequences mentioned in the specification and producing a plant having an	
CC	CC	improved property. Producing a plant having an improved property	
CC	CC	comprises transforming a plant with a recombinant construct comprising a	
CC	CC	promoter region functional in a plant cell operably joined to a	
CC	CC	polynucleotide comprising a coding sequence for a polypeptide associated	
CC	CC	with the property, and growing the transformed plant. The polypeptide is	
CC	CC	useful for improving plant cold tolerance, manipulating growth rate in	
CC	CC	plant cells by modification of the cell cycle pathway, improving plant	
CC	CC	drought tolerance, providing increased resistance to plant disease,	
CC	CC	producing galactomannan for lignin or plant growth regulators), improving	
CC	CC	plant heat tolerance, improving plant tolerance to herbicides, increasing	
CC	CC	the rate of homologous recombination in plants, improving plant tolerance	
CC	CC	to extreme osmotic conditions or to pathogens or pests, improving yield	

CC by modification of photosynthesis, modifying seed oil or protein yield
CC and/or content, improving yield by modification of carbohydrate, nitrogen
CC or phosphorus use and/or uptake, or improving yield by providing improved
CC plant growth and development under at least one stress condition. The
CC polynucleotide and polypeptide may also be used in recombinant DNA
CC constructs, in physical arrays of molecules, as plant breeding markers,
CC or in computer-based storage and analysis systems. The present sequence
CC is a Cotton plant cDNA of the invention. NOTE: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040181830. However only 5585
CC polynucleotide sequences were available, the remaining 5213
CC polynucleotides and all 58798 protein sequences were not present.
XX
SQ Sequence 1211 BP; 366 A; 245 C; 272 G; 328 T; 0 U; 0 Other;

Query Match 38.1%; Score 661; DB 13; Length 1211;
Best Local Similarity 75.1%; Pred. No. 1.9e-130;
Matches 851; Conservative 0; Mismatches 280; Indels 2; Gaps 2;

127 ATGAGAGGGAACAAGTTTGGTGTGATTCGGTACCTCTCATCTTGGCTGCTGCC 186
Db
80 ATAAAAATGGCTAAGCTAAGCTGCGATTCGGTACCTCTCATCCCGCTGCATCGTA 139
Qy
187 TTCACTACACACAGATGCGCTTTTTCGACACAGTCAGAAATATGAGATGCGCTTGCT 246
Db
140 TTCAATATACATCCAGATGCGCTTTTTCACACGCAATCTGATATGAGATGCGCA 199
Qy
247 GCTGCAATTTGAAGCAGAAAATCAATTTGACAGCCAGACAGATTCCTTATGACAGATT 306
Db
200 GAAGCGGTGATGAGAGCACCATTGTACGAGTCAATGCGACTACTAAATGATCAAAAT 259
Qy
307 AGCTGACGACGAGGAAGATAGTTCCTTGAAGACAAATGAGCGTCAGGACCGAG 366
Db
260 AGTATGCAACAGAGCAAAATCGTGGCCCTGGAAGGGGAGGAGCGAAGGATCAGGAG 319
Qy
367 TGCGGACAAATTAAGGCTTGTGTTGAGATCTTGAAGTGAAGGACATAAAAGTTGATC 426
Db
320 TGCCACAAATTAAGACTTGTGTTAAGTCTTGAAGAAAGCTCTTCAAGAGTAAT 379
Qy
427 GGAATGTACAGATGCCAGTGGCTGCTGTAGTTGTTATGGCTTGCATTCGGCTGATTAC 486
Db
380 GACAAACACAGGTACCTGTGCGAGCTGTAGTAATCATGCGGTGCAATCGTGTGATTAT 439
Qy
487 CTGGAAGAAGACTAATAATCCATCTTAAATACCAATATCTGTGCTGCTCAAAATATCCT 546
Db
440 CTGGAAGGACAGTTGCGCTGTCTTAAAGTATCAGAGCTCTGTGCTTTCGAAGTATCCA 499
Qy
547 CTTTTCATATCCAGGATGGATCATCTCTGATGTCAGAAAGCTTGTCTTGCAGCTATGAT 606
Db
500 CTTTGTGATCTCAGGATGGATCAGATCAAGGGTTAAACCTAAGGCTTGTGATTATAG 559
Qy
607 CAGCTGAGCTATATGACGACTTGGATTTTGAACCTGTGCATCTTGAAGACCCAGGGGAG 666
Db
560 GAGCTAATCTATATGACGACATAGTATGATCCCGTGCATACAGACCGGCTTGGGAA 619
Qy
667 CTGATTCATATAACAAATTCAGCTCATTAAGTGGGCAATTCAGATCAGCTGTTTAC 726
Db
620 TTGATCGCATACTACAAGATTCGCCGTCACTACAAATGGGCAATGGATGATGTTTCTAC 679
Qy
727 AAGCAATATTTAGCGGTGTTATCATCTAGAGATGATGGAATTTGCCCTGATTTT 786
Db
680 AAGCAATATTTAGCGGATTAATACTTGAAGATGATGGAATTTGCCCTGATTTT 739
Qy
787 TTTGATCTTTTGGGCTGGAGCTACTCTTTTGAAGACAGCAAGCTGATTTGCTATT 846
Db
740 TTTGATCTTTGAGGAGCTGCTGCCCTTCTCGACAGGACAACTCAATATGCTGT 799
Qy
847 TCTTCTGGAATGACAAATGGAATAATGAGTTTGTCCAGATCTCTTATGCTCTTTACGCG 906
Db
800 TCCTCATGGAATGACAAATGGGCAAGAGAGTTTGTGTATGAGCCATATGCACTTTATGCG 859
Qy
907 TCAGATTTTTCGCCGCTTGGATGGATGCTTTTCAAAATCTACTTGGGACGAATATCT 966

860 TCAGATTTCTTCTGCTGCTTGGCTGGATGCTTTACTTAAATCTGTATGGAATGAGCTATCA 919
Qy
967 CCAAACTGCGCAAAAGCTTACTGGGACGACTGGCTAAGACTCAAAGAGATCAACAGAGGT 1026
Db
920 CCAAAATGGCCAAAAGCTTACTGGGATGACTGGTGGATTAATAAATAATCAATGTT 979
Qy
1027 CGACAATTTATTCGCCAGAGTTTCCAGAACATATAAATTTTGGTGAGCATGGTTCAGT 1086
Db
980 CGACAATTCCTTCTGCTGCTGAAGTATGCAGAACATATAAATTTTGGTGAGCATGGTTCAGC 1039
Qy
1087 TTGGGCGAGTTTTCAGCAGATCTTTCAGCCAAATTAATTAATGATGTCAGGTTGAT 1146
Db
1040 ATCGGCGAGTTTTCGAAAAAATACCTTGCACCTATTAAAGATGATGACGT-CAGGTGAC 1098
Qy
1147 TGAAGTCAATGACCTTACTTACCTTTTGGAGGACAAATTACGTGAAACACATTTGGTGAC 1206
Db
1099 TGAAGTCCA-GGATTTGAGTACTTAACCTGAGGAAAAATATGCCAATACTTTGCGAGAC 1157
Qy
1207 TTGGTTAAAAAGGCTAAGCCCATCCATGGAGCTGCTGCTTTGAAAGCAAT 1259
Db
1158 ATTTGAAAGCGCAAAACCTGCTCCCTGGAACAGATTCTGCTCTTATGGCAAT 1210

RESULT 8

ADD68118
ID ADD68118 standard; DNA; 2180 BP.

AC ADD68118;

XX 15-JAN-2004 (first entry)

XX Glycosyltransferase associated hybrid expression cassette SEQ ID NO:35.

XX galactosyltransferase; transmembrane; glycosyltransferase;
XX glycan processing; ds; hybrid gene expression cassette.

XX Unidentified.

XX WO2003078637-A2.

XX 25-SEP-2003.

XX 18-MAR-2003; 2003WO-IB001626.

XX 19-MAR-2002; 2002US-0365735P.

XX (PLAN-) PLANT RES INT BV.

XX Bakker HAC, Florack DEA, Bosch HJ, Rouwendal GJA;

XX WPI; 2003-902819/82.

XX New nucleic acid encoding a hybrid enzyme comprising a transmembrane
XX region of a plant glycosyltransferase and a catalytic region of a
XX mammalian glycosyltransferase, useful for optimizing glycan processing in
XX organism.

XX Disclosure; SEQ ID NO 35; 139pp; English.

XX The invention relates to a novel nucleic acid encoding a hybrid enzyme
XX comprising a transmembrane region of a plant glycosyltransferase and a
XX catalytic region of a mammalian glycosyltransferase. The nucleic acid
XX molecules are useful for optimising glycan processing in an organism so
XX that a glycoprotein having complex bi-antennary glycans and containing
XX galactose residues on both arms and which are devoid of xylose and fucose
XX can be obtained. The present sequence is used in the exemplification of
XX the invention.

XX Sequence 2180 BP; 648 A; 420 C; 492 G; 520 T; 0 U; 0 Other;

Query Match 37.9%; Score 658.6; DB 10; Length 2180;
Best Local Similarity 75.9%; Pred. No. 1.4e-189;

		Matches	814; Conservative	0; Mismatches	259; Indels	0; Gaps	0;
Qy	409	GGCATAAAAAGTTGATCGGNAATGTACAGATGCCAGTGGCTGCTGTAGTTGTTATGGCT	468				
Db	882	GGGAAAACTGACTTGGGGACCATGGGACAGATGCTGTGGCTGCTGTAGTTATGGCC	941				
Qy	469	TGCAATCGGGCTGATTACCTTGGAAAGACTATTAAATCCATCTTAAAAACCAAAATATCT	528				
Db	942	TGCAGTCGTGAGACTATCTTGAAGGACTGTGTAATCAGTTTAACTATCAAACTCCC	1001				
Qy	529	GTTCGCTCAAAATATCCTCTTTTCATATCCAGGATGGATCACAATCCTGTATGTCAGGAAG	588				
Db	1002	GTTCCTTCAAAATATCCTCTATTATCTCAGGATGGATCTGATCAAGCTGTCAAGAGC	1061				
Qy	589	CTTCCTTTGAGCTATGATCAGCTGACGTATATGTCAGCAGCTTGGATTTTGAACCTGTGCAT	648				
Db	1062	AAGTCATTTGAGCTATAATCAATTAACATATATGTCAGCAGCTTGGATTTTGAACCAAGTGGTC	1121				
Qy	649	ACTGAAAGACACAGGGAGCTGATTGTCATCTACAAAAATTCACGTCAATTACAAAGTGGCA	708				
Db	1122	ACTGAAAGGCTGCGGAACTGACTGCTACTACAAAGATTGCACTCACTCAAGTGGCA	1181				
Qy	709	TTGGATCAGCTGTTTTACAAGCAFAATTTTACCGCTGTTATCATATCAGAGATGATATG	768				
Db	1182	CTGCACCACTGTTTTTACAACAACAAATTTAGTCGAGTGAATTACTAGAAGATGATATG	1241				
Qy	769	GAATTCGCCCTGATTTTGTGACTTTTGGAGCTGGAGCTACTCTTCTTGAAGAGAC	828				
Db	1242	GAATTCCTCCAGACTTCTTTGATTACTTTGAGCTGCAGCTAGTCTCATGGATAGGGAT	1301				
Qy	829	AAGTCGATATTGGCTATTTCTCTTGGAAATGACAATGCAAAATGCAAGTTTGTCCAAGAT	888				
Db	1302	AAACCAATATTGGCTGCTCATCATGGAATGATATGACAGAGCAGTTTGTGCATGAT	1361				
Qy	889	CCTTATGCTCTTTACCGCTCAGATTTTTTCCCGGCTCTTGGATGGATGCTTTCAAAATCT	948				
Db	1362	CCCTATGGCTATACCGATCAGATTTTTTCTGGCCTTGGTGGATGCTCAAGAGATCG	1421				
Qy	949	ACTTGGACGAATATCTCOAAGTGGCCAAAGGCTTACTTGGGACGACTGGCTAAGACTC	1008				
Db	1422	ACTTGGGATGAGTTATCACCAAAGTGGCCAAAGGCTTACTTGGGATGATGGCTGAGACTA	1481				
Qy	1009	AAAGAGATCACAGAGCTGCACAATTTATTCGCCCAAGCTTTCAGAAACATATAATTT	1068				
Db	1482	AGGAAACCAATAAAGCCGCCAATTCATTCGACCGGAAGTCTGTAGAACATCAAAATTT	1541				
Qy	1069	GGTGAGCATGTTCTAGTTTGGGCGAGTTTTTCAAGCAGATATCTTGAGCCAAATTAACATA	1128				
Db	1542	GGTGAAACATGGTCTAGTTTGGGACAGTTTTTTCAGTCAGTATCTGGAACCTATAAAGCTA	1601				
Qy	1129	AATGATGTCAGGTTGATTTGAAGTCAATGGACCTTATAGTTACCTTTTGGAGGACAATTAC	1188				
Db	1602	AACGATGTACGGTTGACTGGAAAGCAAAAGACCTGGGATACCTGCACAGAGGGAAACTAT	1661				
Qy	1189	GTGAAACACTTTGCTGACTTGTGTTAAAGGCTAAAGCCCATCCATGGAGCTGATGCTGTC	1248				
Db	1662	ACCAAGTACTTTTCTGGCTTAGTGAGACAAGCAGCAAAATCAAGTTCTGACCTTTGTC	1721				
Qy	1249	TTGAAACGATTTAAACATAGATGTTGATGTGCGTATTTCAGTACAGAGATCAACTAGACTTT	1308				
Db	1722	TTAAAGGCTCAAAACATAAAGGATGATTTTCGTATCCGTTATAAAGACCAAGTAGAGTTT	1781				
Qy	1309	GAAATATCGCACGGCAATTTTGGCAATTTTGAAGAAATGGAAGGATGTTGACCAAGTGA	1368				
Db	1782	GAACGCATTCAGGGGAAATTTGGTATATTTGAAGAAATGGAAGGATGTTGTCCTCGAACA	1841				
Qy	1369	GCATATAAGGAATAGTAGTTTTCGGTACCACAGCTCCAGAGCTGATTTCTTGTGGC	1428				
Db	1842	GCATATAAGGATAGTAGTTTTCGAAATCCAGAACAAAGACGTGATTTCTTGTGGG	1901				
Qy	1429	CATGATTCGCTTCAACAACTCGGAATTTGAAGATATCTTAAACAAAGATATGATG	1481				
Db	1902	CCAGATTCGTAAATGCAGCTTGGAAATTCGAAATTCCTGATGCGGATCCGCTAG	1954				

RESULT 9	
ADD68117	
ID	ADD68117 standard; DNA; 9240 BP.
XX	
AC	ADD68117;
XX	
DT	15-JAN-2004 (first entry)
XX	
DE	Glycosyltransferase associated triple cassette SEQ ID NO:34.
XX	
KW	galactosyltransferase; transmembrane; glycosyltransferase;
KW	glycan processing; ds; triple cassette.
XX	
OS	Unidentified.
XX	
PN	WO2003078637-A2.
XX	
PD	25-SEP-2003.
XX	
PF	18-MAR-2003; 2003WO-IB001626.
XX	
PR	19-MAR-2002; 2002US-0365735P.
XX	
PA	(PLAN-) PLANT RES INT BV.
XX	
PI	Bakker HAC, Florack DEA, Bosch HJ, Rouwendal GJA;
XX	
DR	WPI; 2003-902819/82.
XX	
PT	New nucleic acid encoding a hybrid enzyme comprising a transmembrane
PT	region of a plant glycosyltransferase and a catalytic region of a
PT	mammalian glycosyltransferase, useful for optimizing glycan processing in
XX	organism.
XX	
PS	Disclosure; SEQ ID NO 34; 139pp; English.
XX	
CC	The invention relates to a novel nucleic acid encoding a hybrid enzyme
CC	comprising a transmembrane region of a plant glycosyltransferase and a
CC	catalytic region of a mammalian glycosyltransferase. The nucleic acid
CC	molecules are useful for optimising glycan processing in an organism so
CC	that a glycoprotein having complex bi-antennary glycans and containing
CC	galactose residues on both arms and which are devoid of xylose and fucose
CC	can be obtained. The present sequence is used in the exemplification of
XX	the invention.
XX	
Seq	Sequence 9240 BP; 2707 A; 1778 C; 2037 G; 2718 T; 0 U; 0 Other;
Query Match 37.9%; Score 658.6; DB 10; Length 9240;	
Best Local Similarity 75.9%; Pred. No. 3.2e-189;	
Matches 814; Conservative 0; Mismatches 259; Indels 0; Gaps 0;	
Qy	409 GGCATAAAAAAGTTGATCGGNAATGTACAGATGCCAGTGGCTGCTGTAGTTGTTATGGCT 468
Db	882 GGGAAAACTGACTTGGGGACCATGGGACAGATGCTGTGGCTGCTGTAGTTATGGCC 941
Qy	469 TGCATCGGGCTGATTACCTTGGAAAAAGACTATTAAATCCATCTTAAAAACCAAAATATCT 528
Db	942 TGCAGTCGTGCAGACTATCTTGAAGGACTGTTAAATCAGTTTAAACATATCAAACTCCC 1001
Qy	529 GTTCGCTCAAAATATCCTCTTTTCATATCCAGGATGGATCACAATCCTGTATGTCAGGAAG 588
Db	1002 GTTCCTTCAAAATATCCTCTATTATCTCAGGATGGATCTGATCAAGCTGTCAAGAGC 1061
Qy	589 CTTCCTTTGAGCTATGATCAGCTGAGCTATATGTCAGCAGCTTGGATTTTGAACCTGTGCAT 648
Db	1062 AAGTCATTTGAGCTATAATCAATTAACATATATGTCAGCAGCTTGGATTTTGAACCAAGTGGTC 1121
Qy	649 ACTGAAAGACACAGGGAGCTGATTGTCATCTACAAAAATTCACGTCAATTACAAAGTGGCA 708
Db	1122 ACTGAAAGGCTGCGGAACTGACTGCTACTACAAAGATTGCACTCACTCAAGTGGCA 1181
Qy	709 TTGGATCAGCTGTTTTACAAGCAFAATTTTACCGCTGTTATCATATCAGAGATGATATG 768
Db	1182 CTGCACCACTGTTTTTACAACAACAAATTTAGTCGAGTGAATTACTAGAAGATGATATG 1241
Qy	769 GAATTCGCCCTGATTTTGTGACTTTTGGAGCTGGAGCTACTCTTCTTGAAGAGAC 828
Db	1242 GAATTCCTCCAGACTTCTTTGATTACTTTGAGCTGCAGCTAGTCTCATGGATAGGGAT 1301
Qy	829 AAGTCGATATTGGCTATTTCTCTTGGAAATGACAATGCAAAATGCAAGTTTGTCCAAGAT 888
Db	1302 AAACCAATATTGGCTGCTCATCATGGAATGATATGACAGAGCAGTTTGTGCATGAT 1361
Qy	889 CCTTATGCTCTTTACCGCTCAGATTTTTTCCCGGCTCTTGGATGGATGCTTTCAAAATCT 948
Db	1362 CCCTATGGCTATACCGATCAGATTTTTTCTGGCCTTGGTGGATGCTCAAGAGATCG 1421
Qy	949 ACTTGGACGAATATCTCOAAGTGGCCAAAGGCTTACTTGGGACGACTGGCTAAGACTC 1008
Db	1422 ACTTGGGATGAGTTATCACCAAAGTGGCCAAAGGCTTACTTGGGATGATGGCTGAGACTA 1481
Qy	1009 AAAGAGATCACAGAGCTGCACAATTTATTCGCCCAAGCTTTCAGAAACATATAATTT 1068
Db	1482 AGGAAACCAATAAAGCCGCCAATTCATTCGACCGGAAGTCTGTAGAACATCAAAATTT 1541
Qy	1069 GGTGAGCATGTTCTAGTTTGGGCGAGTTTTTCAAGCAGATATCTTGAGCCAAATTAACATA 1128
Db	1542 GGTGAAACATGGTCTAGTTTGGGACAGTTTTTTCAGTCAGTATCTGGAACCTATAAAGCTA 1601
Qy	1129 AATGATGTCAGGTTGATTTGAAGTCAATGGACCTTATAGTTACCTTTTGGAGGACAATTAC 1188
Db	1602 AACGATGTACGGTTGACTGGAAAGCAAAAGACCTGGGATACCTGCACAGAGGGAAACTAT 1661
Qy	1189 GTGAAACACTTTGCTGACTTGTGTTAAAGGCTAAAGCCCATCCATGGAGCTGATGCTGTC 1248
Db	1662 ACCAAGTACTTTTCTGGCTTAGTGAGACAAGCAGCAAAATCAAGTTCTGACCTTTGTC 1721
Qy	1249 TTGAAACGATTTAAACATAGATGTTGATGTGCGTATTTCAGTACAGAGATCAACTAGACTTT 1308
Db	1722 TTTAAAGGCTCAAAACATAAAGGATGATTTTCGTATCCGTTATAAAGACCAAGTAGAGTTT 1781
Qy	1309 GAAATATCGCACGGCAATTTTGGCAATTTTGAAGAAATGGAAGGATGTTGACCAAGTGA 1368
Db	1782 GAACGCATTCAGGGGAAATTTGGTATATTTGAAGAAATGGAAGGATGTTGTCCTCGAACA 1841
Qy	1369 GCATATAAGGAATAGTAGTTTTCGGTACCACAGCTCCAGAGCTGATTTCTTGTGGC 1428
Db	1842 GCATATAAGGATAGTAGTTTTCGAAATCCAGAACAAAGACGTGATTTCTTGTGGG 1901
Qy	1429 CATGATTCGCTTCAACAACTCGGAATTTGAAGATATCTTAAACAAAGATATGATG 1481
Db	1902 CCAGATTCGTAAATGCAGCTTGGAAATTCGAAATTCCTGATGCGGATCCGCTAG 1954

QY 709 TTGGATCAGCTGTTTACAGCATAATTTTAGCCGTGTATCATACATAGAGATGATG 768
DB 1182 CTGGACCAAGTGTGTTTACAAACACAAATTTAGTCGAGTGATATACATAGAGATGATG 1241
QY 769 GAAATGCCCCCTGATTTTGTGACTTTTGTGAGCTGGAGCTACTCTTCTTGAAGAGAC 828
DB 1242 GAAATGCTCCAGACTCTTTTGATTTACTTTGAGGCTGCAGCTAGTCTCATGGATGGAT 1301
QY 829 AAGTCGATATGCGTATTTCTTCTGGAATGACAAATGGACAAATGCAAGTGTGCTCAAGAT 888
DB 1302 AAAACCAATATGCTGCTTCATCATGGAATGATAATGGACAGAGAGTGTGTCATGAT 1361
QY 889 CTTATGCTCTTTACCGCTCAGATTTTTCCTCCGCTCTTGGATGGATGCTTTCAAAATCT 948
DB 1362 CCTATGGCTATACCGATCAGATTTTCTTGGCTTGGTGGATGCTCAAGATCG 1421
QY 949 ACTTGGGACGAATATCTTCAAGTGGCCAAAGGCTTACTGGACGACTGGCTGAAGCTC 1008
DB 1422 ACTTGGGATGATTTATCACCAGAGTGGCCAAAGGCTTACTGGATGATGGCTGAGACTA 1481
QY 1009 AAAGGAATACAGAGTGCAGATTTTATTCGCCAGAGTTCGACGAACATATATTTT 1068
DB 1482 AAGGAAAACCAATAAGGCGGCAATTTCAATCGACGGAAGTCTGTAGAACAATCAATTTT 1541
QY 1069 GGTGAGCATGTTCTAGTTTGGGCGAGTTTTCAGACGAGTATCTTGAGCCCAATTAACATA 1128
DB 1542 GGTGAACATGGTCTAGTTTGGGACAGTTTTCAGTCTAGTATCTGGAACTTATAAGCTA 1601
QY 1129 AATGATGTCAGGTTGATTTGAAGTCAATGACACTTACTTACCTTTTGGAGCAATATAC 1188
DB 1602 AACGATGTGACGTTGACTGGAAAGCAAGGACCTGGATACCTGACAGAGGGAACATAT 1661
QY 1189 GTGAAACACTTTGTTGACTGTTGTTAAAGGCTAAGCCATCCATGGAGCTGATGCTGTC 1248
DB 1662 ACCAAGTACTTTTCTGGCTTAGTGAGACAGCACGACCAATTTCAAGGTTCTGACCTGTC 1721
QY 1249 TTGAAGCATTTAAACATAGATGTCATGTCGTATTTCAAGTACAGAGATCAACTAGACTTT 1308
DB 1722 TTAAGGCTCAAAACATAAAGATGATGTTGATTCGGTATTAAGACNAGTAGAGTTT 1781
QY 1309 GAAATATCGCAGCGCAATTTGGCAATTTTGAAGATGGAAGGATGGTGTACCAAGTGA 1368
DB 1782 GAACCAATGCAAGGGGAAATTTGATATTTTGAAGATGGAAGATGGTGTGCTCGAACA 1841
QY 1369 GCATATAAGGAATAGTAGTTTTCGGTACCAAGCTCCAGACGTGATTTCTGTTGGC 1428
DB 1842 GCATATAAGGAGTAGTGTGTTTCCGAATCCAGACAACAGACGTGATTTCTGTTGGG 1901
QY 1429 CATGATTCGCTTCAACACTCGGAATGGAATGATCTTAAACAAGATATGATTG 1481
DB 1902 CCAGATTTCTGTAATGAGCTTGGAAATTCGAATTCGATGCGGATCCGCTAG 1954

RESULT 10
ADD68120
ID ADD68120 standard; DNA; 1319 BP..
XX AC
XX ADD68120;
XX
DT 15-JAN-2004 (first entry)
XX
DE Glycosyltransferase associated hybrid gene fusion SEQ ID NO:37.
XX
KW galactosyltransferase; transmembrane; glycosyltransferase;
KW glycan processing; ds; hybrid.
XX
OS Unidentified.
XX
XX WO2003078637-A2.
XX
PD 25-SEP-2003.
XX
PF 18-MAR-2003; 2003WO-1B001626.

XX 19-MAR-2002; 2002US-0365735P.
PR (PLAN-) PLANT RES INT BV.
XX Bakker HAC, Florack DBA, Bosch HJ, Rouwendal GJA;
XX WPI; 2003-902819/82.
DR
XX New nucleic acid encoding a hybrid enzyme comprising a transmembrane
PT region of a plant glycosyltransferase and a catalytic region of a
PT mammalian glycosyltransferase, useful for optimizing glycan processing in
PT organism.
XX
PS Disclosure; SEQ ID NO 37; 139pp; English.
XX
CC The invention relates to a novel nucleic acid encoding a hybrid enzyme
CC comprising a transmembrane region of a plant glycosyltransferase and a
CC catalytic region of a mammalian glycosyltransferase. The nucleic acid
CC molecules are useful for optimising glycan processing in the organism so
CC that a glycoprotein having complex bi-antennary glycans and containing
CC galactose residues on both arms and which are devoid of xylose and fucose
CC can be obtained. The present sequence is used in the exemplification of
CC the invention.
XX
SQ Sequence 1319 BP; 380 A; 249 C; 328 G; 362 T; 0 U; 0 Other;
Query Match 37.9%; Score 658; DB 10; Length 1319;
Best Local Similarity 76.1%; Pred. No. 1.6e-189;
Matches 811; Conservative 0; Mismatches 255; Indels 0; Gaps 0;
QY 409 GGCATAAAAAAGTTGATCGGAAATGTACAGATGCCAGTGGCTGTGTATGTTATGGCT 468
DB 252 GGGAAACTGACTTTGGGACCATGGACAGATGCTGTGGCTGTGTATGTTATGGCC 311
QY 469 TGCATCGGGCTGATTTACCTGGAAAAGACTATTAATCCATCTTAAATACCAATATCT 528
DB 312 TGCAGTCGTGACGACTATCTTTGAAAGGACTGTAAATCAGTTTAAACATATCAAACTCCC 371
QY 529 GTTCGTCAAAATATCTCTTTTTCATATCCAGGATGGATCAGATCTCTGATGTCAGGAAG 588
DB 372 GTTGCTTCAAAATATCTCTTATATCTCAGAGTGGATCTGATCAAGCTGTCAAGAGC 431
QY 589 CTGTGTTGAGCTATGATCAGCTGACGTATATGACGACTTGGATTTTGAACCTGTGCAT 648
DB 432 AAGTCATTTGAGCTATATCAATTAACATATATGACGACTTGGATTTTGAACCACTGGTC 491
QY 649 ACTGAAAGACAGGGGAGCTGATTCATCTACTACAAATTTGACGTCATTAAGTGGGCA 708
DB 492 ACTGAAAGGCTTGGCGAACTGACTGCTGCTACTACAAAGATTGACAGTCTCAAGTGGGCA 551
QY 709 TTGGATCAGCTGTTTACAGCATAATTTTAGCCGTGTATCATACATAGAGATGATG 768
DB 552 CTGGACCAAGTGTGTTTACAAACACAAATTTAGTCGAGTGATATACATAGAGATGATG 611
QY 769 GAAATGCCCCCTGATTTTGTGACTTTTGTGAGGCTGGAGCTACTCTTCTTGAAGAGAC 828
DB 612 GAAATGCTCCAGACTCTTTGATTTACTTTGAGGCTGCAGCTAGTCTCATGATAGGAT 671
QY 829 AAGTCGATATGCGTATTTCTTCTGGAATGACAAATGGACAAATGCAAGTGTGCTCAAGAT 888
DB 672 AAAACCAATATGCGTCTTCATCATGGAATGATAATGGACAGAGAGTGTGTCATGAT 731
QY 889 CTTATGCTCTTTACCGCTCAGATTTTTCCTCCGCTCTTGGATGGATGCTTTCAAAATCT 948
DB 732 CCTATGCGCTATACCGATCAGATTTTTCCTGGCTTGGTGGATGCTCAAGAGATCG 791
QY 949 ACTTGGGACGAATTTATCTCCAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTGAAGCTC 1008
DB 792 ACTTGGGATGATTTATCACCAGTGGCCAAAGGCTTACTGGATGATGTTGGCTGAGACTA 851
QY 1009 AAAGAGAAATCACAGAGGTCGCAATTTTATTCGCCCAAGAGTTTTCGAGAACATATATTTT 1068

Db 852 AAGGAACACATTAAGCGCCCAATTTCATTCGACCGGAAGTCTGTAGAACATACAAATTTT 911
Qy 1069 GGTGAGCATGTTCTAGTTGGGCGAGTTTTCGAAGCAGATCTTTGAGCCAAATTAACATA 1128
Db 912 GGTGAACATGGGTCTAGTTTGGGACAGTTTTCAGTCAGTATCTGGAACCTATAAAGCTA 971
Qy 1129 AATGATGTCGAGGTGATTTGGAAGTCAATGACCTTAGTTACCTTTTGGAGGACAAATTAC 1188
Db 972 AACGATGTGACGGTGTGACTGGAAGCAAGAGACCTGGGATACCTTGAACAGAGGGAACTAT 1031
Qy 1189 GTGAACACATTTGTGACTTGGTTTAAAAAGGCTAAGCCCATCCATCGAGCTGATGCTGTC 1248
Db 1032 ACCAAGTACTTTTCTGGCTTAGTGAGACAGCACCAATTCAAGTCTTGACCTTGTGTC 1091
Qy 1249 TTGAAGCATTTACATAGATGTTGATGCGGTATTCAGTACAGAGATCAACTAGACTTT 1308
Db 1092 TTAAGGCTCAAAACATATAAGGATGATGTCGTATCCGGTATAAAGACCAAGTAGAGTTT 1151
Qy 1309 GAAATATCCGACCGCAATTTGGCATTTTTCMAAGATGGAAGGATGGGTACCAACGTCGA 1368
Db 1152 GAACGCAATTCAGGGGAATTTGGTATATTTGAAGATGGAAGGATGGTGGCTTCGAACA 1211
Qy 1369 GCATATAAGGAATAGTAGTTTTCGGTACCAACGTCAGACGCTGATTTCCCTTGTGGC 1428
Db 1212 GCATATAAGGATAGTGTGTTTTCGAATCCAGACACAGACGCTGATTTCCCTGTTGG 1271
Qy 1429 CATGATTGCTTCAACACTCGGAATTTGAAGATACCTTAACAAAGAT 1474
Db 1272 CCAGATTCGTGAATGCAGCTTGGAAATTCGAAATTCCTGATCGCGAT 1317

RESULT 11

ADD68136
ID ADD68136 standard; DNA; 1119 BP.
XX AC ADD68136;
XX AC
XX AC
DT 15-JAN-2004 (first entry)
XX DE Plant GntI/mouse signal peptide hybrid gene SEQ ID NO:53.
XX DE galactosyltransferase; transmembrane; glycosyltransferase;
XX KW glycan processing; ds; gene; plant; mouse.
XX OS Chimeric.
OS Arabidopsis thaliana.
OS Mus sp.
XX FH Key Location/Qualifiers
FT CDS 1..1119
FT /*tag= a
XX WO2003078637-A2.
XX 25-SEP-2003.
XX 18-MAR-2003; 2003WO-IB001626.
XX 19-MAR-2002; 2002US-0365735P.
XX (PLAN-) PLANT RES INT BV.
XX Bakker HAC, Florack DEA, Bosch HJ, Rouwendal GJA;
XX WPI; 2003-902819/82.
XX P-PSDB; ADD68137.
XX PT New nucleic acid encoding a hybrid enzyme comprising a transmembrane
PT region of a plant glycosyltransferase and a catalytic region of a
PT mammalian glycosyltransferase, useful for optimizing glycan processing in
XX organism.
XX PS Disclosure; SEQ ID NO 53; 139pp; English.

XX The invention relates to a novel nucleic acid encoding a hybrid enzyme
CC comprising a transmembrane region of a plant glycosyltransferase and a
CC catalytic region of a mammalian glycosyltransferase. The nucleic acid
CC molecules are useful for optimizing glycan processing in an organism so
CC that a glycoprotein having complex bi-antennary glycans and containing
CC galactose residues on both arms and which are devoid of xylose and fucose
CC can be obtained. The present sequence is used in the exemplification of
CC the invention.
XX
Qy Sequence 1119 BP; 326 A; 203 C; 278 G; 312 T; 0 U; 0 Other;
Query Match 37.8%; Score 656.6; DB 10; Length 1119;
Best Local Similarity 77.0%; Pred. No. 4e-189;
Matches 800; Conservative 0; Mismatches 239; Indels 0; Gaps 0;
Qy 425 TCAGAAATGTACAGATGCCAGTGGCTGTGTAGTTGTATGGCTTGCATCGGCGCTGATT 484
Db 65 TCGACATGGGACAGATGCCGTGGCTGTGTAGTTGTATGGCTTGCAGTCTGTCAGACT 124
Qy 485 ACCTGAAAGACTATTAATTCATCTTAAATACCAATATATCTGTTGCGTCAAAATATC 544
Db 125 ATCTTGAAAGGACTGTAAATCAGTTTAAACATATCAAACTCCCGTTGCTTCAAAATATC 184
Qy 545 CTCTTTTCATATCCAGGATGGATCACATCCTGATCTCAGGAGCTTGTCTTGAGCTATG 604
Db 185 CTCTATTATATCTCAGGATGGATCTGATCAAGCTGTCAAGAGCAAGTCAATTGAGCTATA 244
Qy 605 ATCAGCTGACGTATATGCAGCACTTGGATTTTGAAGCTGTGTCATCTGAAAGACCAAGGG 664
Db 245 ATCAATTAACATATATGCAGCACTTGGATTTTGAACAGTGGTCACTGAAAGCGCTGGCG 304
Qy 665 AGCTGATGCACTACATAAAATTCACGTCATTAAGTGGGCAATTTGATCAGCTGTTTT 724
Db 305 AACTGCTGCTACTACAAGATTGACGTCACCTACAAGTGGGCACTGGAACAGTTGTTTT 364
Qy 725 ACAAGCATATTTTAGCCGCTGTATCATATAGAGATGATATGGAATTTGCCCTGATT 784
Db 365 ACAACACAAATTTAGTCAGTGATTTACTAGAAGATGATATGGAATTTGCTCCAGACT 424
Qy 785 TTTTTCAGCTTTTGTAGGCTGGAGCTACTCTTCTGACAGAGACAAAGTCAATTTAGCTA 844
Db 425 TCTTGTATTTAGGCTGGAGCTAGTCTCATGATAGGATATAACCAATTTATGGCTG 484
Qy 845 TTTCTTCTTGGAAATGACAAATGACAAATGCAAGTTTGTCCAAATCTCTATGCTCTTTACC 904
Db 485 CTTTCATCATGAATGATAATGACAGAAAGCAGTTTGTGATGATCCCTATGCGCTATACC 544
Qy 905 GCTCAGATTTTTCCTGGCTTGGATGATGCTTTTCAAAATCTACTTTGGGAGCAATAT 964
Db 545 GATCAGATTTTTCCTGGCTTGGGATGATGCTCAAGAGATCGACTTTGGGATGAGTTAT 604
Qy 965 CTCCTAAGTGGCCAAAGGCTTACTGGGACGACTGGCTTAAGACTCAAGAGAAATCACAGAG 1024
Db 605 CACCAAGTGGCCAAAGGCTTACTGGGATGATTTGGCTGAGACTTAAAGGAAACCAATAAG 664
Qy 1025 GTCGACAAATTTATTCGCCAGAAAGTTTGCAGAAACATATATTTTGTGAGCATGTTCTA 1084
Db 665 GCGCCAAATTCATTTCGACCGGAAGTCTGTAGAACATACAAATTTTGGTGAACAGGCTTA 724
Qy 1085 GTTTGGGCGAGTTTTCAGGCAAGTCTGAGCCAAATTAACCTAAATGATGTCAGGTTG 1144
Db 725 GTTTGGGACAGTTTTCAGTCAAGTATCTGGAACCTATAAGCTTAAACGATGTGACGGTTG 784
Qy 1145 ATTGGAAGTCAATGGACCTTAGTTTACCTTTTGGAGCAATTAACGTAACACCTTTGGTG 1204
Db 785 ACTGGAAGCAAGGACCTGGGATACCTGACAGAGGGAACATATACCAAGTACTTTTCTG 844
Qy 1205 ACTTGGTTAAAAAGGCTAAGCCCATCCATGGAGCTGATGCTGTCTTGAAGCATTTAACA 1264
Db 845 GCTTAGTGAGCAAGCAAGCAGCAATTTCAAGGTTCTGACCTTGTCTTAAAGGCTCAAAACA 904
Qy 1265 TAGATGCTGATGTCGTATTTTCAGTACAGAGATCAACTAGACTTTTGAANAATATCGCAGCGC 1324

Db 905 TAAAGGATGTTTCGATCGGTATTAAGACCAAGTAGAGTTTGAACGCATTGCAGGG 964
Qy 1325 AATTGGCATTTTGAAGATGGAAGATGGTGTACACAGTGCAGCATATAAGGATAG 1384
Db 965 AATTGGTATATTGAAGATGGAAGATGGTGTGCCTCGAACAGCATATAAGAGGATAG 1024
Qy 1385 TAGTTTTCGGTACCAACAGTCCAGACGTGTATTCCTTTGTTGGCCATGATTCGCTTCAAC 1444
Db 1025 TGGTGTTCGATCCAGACAAAGACGTGTATTCCTTGGTGGCCAGATTCCTGTAATGC 1084
Qy 1445 AACTCGGAATTGAAGATAC 1463
Db 1085 AGCTTGAATTCGAAATTC 1103

RESULT 12
ADT20185/c
ID ADT20185 standard; cDNA; 1898 BP.

XX AC ADT20185;
XX 13-JAN-2005 (first entry)
XX Plant cDNA, Seq ID 5511.
XX Plant; ss; gene; transgenic; cold tolerance; growth rate;
KW drought tolerance; disease resistance; galactomannan production;
KW plant growth regulator; heat tolerance; herbicide tolerance;
KW lignin production; extreme osmotic condition tolerance;
KW pathogens resistance; pest resistance; yield improvement; seed oil yield;
KW seed protein yield.
XX Viridiplantae.

OS US2004216190-A1.
XX 28-OCT-2004.
XX 18-DEC-2003; 2003US-00739930.
XX 28-APR-2003; 2003US-00424599.
XX 28-APR-2003; 2003US-00425115.
XX (KOVA/) KOVALIC D K.
XX Kovalic DK;
XX WPI; 2004-757369/74.
XX New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
PT biological characteristics.
XX Claim 1; SEQ ID NO 5511; 14pp; English.

XX The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomannan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving

CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant cDNA
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040216190.
XX
SQ Sequence 1898 BP; 470 A; 504 C; 405 G; 518 T; 0 U; 1 Other;

Query Match 33.5%; Score 581.2; DB 13; Length 1898;
Best Local Similarity 66.4%; Pred. No. 5.6e-166;
Matches 887; Conservative 0; Mismatches 429; Indels 19; Gaps 3;

Qy 147 CTGTGATTTCGGTACCTCTCATCTTGGCTGTGTGCTTCATCTACACACAGATGCG 206
Db 1778 CTGGGACATCCGCTCTCTCTGTCGGCGCGCTGCTTCATCTACATCCAGTGG 1719
Qy 207 GCTTTTTCGACACAGTGCAGATATGCAGATCGCTTGTCTGCTGCTCAATTTGAAGC----- 260
Db 1718 GCTGTCTCCACCCAGTCTCTCTCTTCTTAATTCCTCCTTCCAGACGCGGACG 1659
Qy 261 -----AGAAATCATTTGTACAGCCACAGATGCTTATTTGACCCAGATGAGCCT 311
Db 1658 CCTCGCGAAGCAGAGAAATCGGAGAGCCAGCTGCACGCCATGANGCAGCCAGGTGAGCGC 1599
Qy 312 GCAGCAGGAGAGATAGTGTCTTTGAAGACAAATGAAGCGTCAGGACCCAGGAGTGGCG 371
Db 1598 CCAGCAGGAGAGAAATCGCGCCTTTGAAGATGAGGTAGCCNAGACGAGGAGCGTGT 1539
Qy 372 ACAATTAAAGGCTCTTGTTCAGGATCTTTGAAGTAAGGGGCATATAAAAGTTGATCGGAAA 431
Db 1538 GCAGTCAAGATTTAATCCAAGATCTTGAAGAGGAGTTTACAGACCTTTAAACAAACA 1479
Qy 432 TGTACAGATGCCAGTGGCTGCTGTAGTTGTTATGGCTTGCATTCGGGCTGATTACTGGA 491
Db 1478 GAACTGGTCTCTGTCGCCGCTGTGTCTAATGGCTTGCATTCGACCATCGACGACTATTTCGA 1419
Qy 492 AAAGACTATTAAATCCCATCTTAAATACCAATATCTGTGCTCAAAATATCCCTCTTTT 551
Db 1418 GAGGACAGTGGAAATCTATCTCAAGTACCAAGACAGTGTGCTTCAAGTTTCCACTATT 1359
Qy 552 CATATCCAGGATGGATCACATCTCTGTGTGAGGAAGCTTGTCTTTGAGCTATGATCAGCT 611
Db 1358 TATATCAGAGATGGAACAAATGGAGAGTAAAGAAAGAGCTTTGAGTTACAATCAAT 1299
Qy 612 GACGTATATGACGCA-CTTGAATTTTGAACCTGTGCATCTGAAGACAGCGGGAGCTGA 670
Db 1298 AACATTTATGAGCATGTGGATCTTGAGCCTGTGCGCACTGTAAGAGACAGGAGAGAAAG 1239
Qy 671 TTGCATCTACAAATTCGACGTCTATCAAGTGGGCATTTGCATCAGCTGTTTTCACAGC 730
Db 1238 TTGCATATTCAAGATAGTCAACACTATAATGGGCTTGGATGAGCTATTCAATAGC 1179
Qy 731 ATAATTTAGCGCTTTATCTACTAGAGATGATGAAATTCGCCCTCGATTTTTTTG 790
Db 1178 ATGATTTTCGTGAGTATCATCTTGGAGATGACATGGAGATCGCCCGAGATTTCTTTG 1119
Qy 791 ACTTTTTTGGGCTGAGCTACTCTTCTTGAAGAGACAGTTCGATTTATGGCTATTCTT 850

Db 1118 ACTACTTTGAGCTGCAGCGAATATTCTTGACACTGCAGACACAATAATGGCTGTTTCAT 1059
Qy 851 CTTGGAATGCAATGGGCAATATGAGCTTTGTCGAAGATCCTTATGCTCTTTACCGCTCAG 910
Db 1058 CTTGGAATGATAATGGGCAAAAGCAGTTGCTTTATGACCCCAAGGCTCTTTACCGTTCCG 999
Qy 911 ATTTTTCCTCCGGTCTTTGGATGGATGCTTTTCAAAATCTACTTGGGACGAATTTCTCCAA 970
Db 998 ACTTCTTTCGGGCTTTGGATGGATGCTTAAGAAATCAATGATGAGACTGTCGCCAA 939
Qy 971 AGTGCCCAAGGCTTACTGGGACGACTGGCTAAGACTCAAGAGAAATCACAGAGGTCGAC 1030
Db 938 AGTGCCCAAGCTTATTGGGATCACTTGGGTGAGGCTAAAGAGGTACACAGAGATCGC 879
Qy 1031 AATTATTTCGCCAAGATTTGCAAGACATATAAATTTTGGTGAGCATGGTTCATGTTGG 1090
Db 878 AGTTTATTTCGCCCAAGAGTATGCAAGGACATACAACTTTGGCGAGCATGGATCAAGCATGG 819
Qy 1091 GGCAGTTTTTCAAGCAGTATCTTCGAGCAATTAACACTAAATGATGTCAGGTTGATTGGA 1150
Db 818 GACAATTTCTTGATCAATATCTTGAACCTATCAAGTTAAATGATGCTCATATTGACTGGA 759
Qy 1151 AGTCAATGGACCTTAGTTACCTTTTGGAGGACATTAAGTCAACACTTTTGGTACTTGG 1210
Db 758 ATTCGAGGACCTGAGCTACCTCAAGGAGGACAAAGTTTTTGACCAAAATTCGGGAAAGACG 699
Qy 1211 TTAAGGAGGTAAGCCATCATGAGAGTATGCTGCTTTGAAAGCATTTAAACATAGATG 1270
Db 698 TGGCTAGCGGCACACCTGTGCTATGGATCCGATGCTCTGTTGAAGGCCCAACAATCTGGATG 639
Qy 1271 GTGATGTGCTGATTCAGTACAGAGATCAACTAGACTTTTGAATAATATCGACGGCAATTTG 1330
Db 638 TGGACGTAAGATTCAGTATGCAATCAGGCGGACTTTGAGCGTATAGCTCGGCGATTTG 579
Qy 1331 GCATTTTGAAGATGGAAGGATGGTGTACACGCTGCAGCATATAAAGGAATAGTAGTTT 1390
Db 578 GAATATTTTGAAGAGTGGAGGACGGTGTCCACGGCGGCTTACAAAGGCGTGGTGGTGT 519
Qy 1391 TCCGGTACCA---AACGTCCAGAGTGTATTCCTTTGTTGGCATGATTCGCTTCAACAAC 1447
Db 518 TCCGGTCAAGGGCAGCGGAGAGCGGATATACCTGTGGTCCGCGTCTCTTCGCCCAAC 459
Qy 1448 TCGGAATTTGAAGATA 1462
Db 458 TCGGGGTTAGCTTA 444

RESULT 13
ID AEB65577 standard; DNA; 896 BP.
AC AEB65577;
XX
DT 22-SEP-2005 (first entry)
XX
XX Rice genome derived DNA sequence, SEQ ID 722.
DE
XX transcription; gene regulation; transgenic plant; RNA interference;
KW transformation; antibody; ds.
XX
XX Oryza sp.
XX
PN JP2005185101-A.
XX
XX 14-JUL-2005.
XX
XX 11-DEC-2002; 2002JP-00383870.
XX
XX 30-MAY-2002; 2002JP-00203269.
XX
XX (DOKU-) DOKURITSU GYOSHI HOJIN NOGYO SEIBUTSU SH.
PA (SEIB-) SEIBUTSUKI TOKUTEI SANGYO GIJUTSU.
PA (DOKU-) DOKURITSU GYOSHI HOJIN RIKAGAKU KENKYUSH.

PA (KOKU-) ZH KOKUSAI KAGAKU SHINKO ZAIDAN.
XX
PI Kikuchi H, Hayashizaki Y, Otono Y, Matsubara K, Murakami K;
PI Kishimoto N, Sato K, Nagata T, Kawakami N, Yazaki J, Ishikawa M;
PI Doi K, Kawai J;
XX
DR WPI; 2005-566181/58.
XX
XX Novel DNA encoding transcription factor, derived from rice plant, useful
PT for obtaining transcriptional-regulatory regions in plant and for
PT producing modified plant.
XX
PS Claim 1; SEQ ID NO 722; 2928pp; Japanese.
XX
XX The invention relates to a novel DNA sequence encoding a transcription
CC factor derived from a plant. The invention further comprises antisense
CC RNA sequences, ribozyme activity RNA, RNAi sequences, a vector,
CC transformed plant cells, antibodies and proteins, all related to the
CC novel plant DNA sequences of the invention. The novel DNA is preferably
CC derived from a rice-genome database. The invention further provides a
CC method for determining the transcription regulatory regions of the rice
CC genome. The novel DNA is useful for controlling the expression of a gene
CC in a plant and for producing a modified plant with desired and different
CC characteristics. The plant DNA and method enables the acquisition of many
CC transcriptional-regulatory regions. This polynucleotide represents a DNA
CC sequence taken from a rice genome clone library for use in the invention.
CC Note: This sequence is not shown in the specification. It has been
CC retrieved from a sequence listing in electronic format from the Japanese
CC Patent Office. The invention claims DNA sequences of SEQ ID Nos 1 to
CC 28469 and encoded protein sequences of SEQ ID Nos 28470 to 56791,
CC however, the sequence listing only provided the DNA sequences of SEQ ID
CC Nos 1 to 3032.
XX
SQ Sequence 896 BP; 256 A; 184 C; 220 G; 236 T; 0 U; 0 Other;
Query Match 24.5%; Score 425.8; DB 14; Length 896;
Best Local Similarity 73.6%; Pred. No. 1e-118;
Matches 557; Conservative 0; Mismatches 197; Indels 3; Gaps 1;
Qy 695 ATTCAAGTGGGCAATGGATCAGCTGTTTACAAGCATAAATTTAGCGCTTTATCATAC 754
Db 1 ATTCAAAATGGGCTTTGGATGAGCTATTATCAACACAACTTTGGCGGAGTAATAATTC 60
Qy 755 TAGAAGATGATGGAATTTGCCCTGATTTTTCGACTTTTGGAGCTGGAGCTACTC 814
Db 61 TGAAGATGATATGGAGATGCGCCCTGATTTTTCGACTACTTCGAGGCTGCAGCTAAAT 120
Qy 815 TTCTTGACAGACAAAGTGGATTTATGGCTATTCTTCTTGGATGACAAATGCAATGC 874
Db 121 TACTTGATTAACGTAAGACAAATCATGCGCGTTCCTTCTGGATGACATGGACAAAGC 180
Qy 875 AGTTTCTCAAGATCCTTATGCTTTTACCGCTCAGATTTTTCCTCGGCTTTGGATGGA 934
Db 181 AATTTGTTTACGACCCAAAAGGCTTTTACCGGTGCGATTTCTTCTCGGACTTGGATGA 240
Qy 935 TGCTTTCAAAATCTACTTGGGACGATTTATCTCAAGTGGCCCAAGGCTTACTGGGACG 994
Db 241 TGTTAAAGGCTTACATGATTTGAGCTGTACCTAAGTGGCCCAAGCTTATTTGGATG 300
Qy 995 ACTGGCTAAGACTCAAGAGAAATCACAGAGGTCGACAAATTTATTCGCCCAAGAGTTGCA 1054
Db 301 ACTGGGTGAGGCTTAAAGAGGTACACAGAGACCGGCAATTTATTCGCCCGGAGTCTGCA 360
Qy 1055 GAACATATAATTTTGGTGAGCATGTTCTAGTTTGGGCGAGTTTTTCAAGCAGTATCTTG 1114
Db 361 GAACATATAATTTTGGGAAACATGGGTCAAGCATGGGCAATTTCTTTCAGACAGTACTTG 420
Qy 1115 AGCAATTAACCTAATGATGTCAGGTTGATTTGGAAGTCAATGGACCTTAGTACTCTT 1174
Db 421 AACCAATCAAGCTTAAATGATGCCCATATCAAGTGGAAATTTCTGAGGACCTTACCTCA 480
Qy 1175 TGAGGACAAATTTACGTGAAACACTTTGGTGACTTGGTTAAAGAGGCTAGGCCCATCATG 1234

Db 481 AGGAGCAAGTCTCTGATCCATTTGGGAAGACGTCGCTAGTGCACCCCTCTCCATG 540
Qy 1235 GAGCTGATGCTGCTTGAAGCAATTAACATAGATGGTATGCTGATTCAGTACAGAG 1294
Db 541 GATCTGATGCGCGTGTGAAGCCACACATATGGAATGGATGGATGTAAGGATCCAGTATAACG 600
Qy 1295 ATCAACTAGACTTTGMAAATATCGCACGGCAATTTGGCATTTTGAAGAATGGAAGGATG 1354
Db 601 ATCAGGAAGACTTTGAGCGGATAGCTCTCAATTTGGAATATTTGAAGAATGGAAGGATG 660
Qy 1355 GTGTACCACTGTCAGCATATATAAGGAATAGTAGTTTTCGGTACCA---AACGTCACAGAC 1411
Db 661 GCATTCGAAGAACAGCTTATAAAGGAGTAGTGTCTTCGGGTACAAAGAGTAGTGAAGAC 720
Qy 1412 GTGTATTCCTTTGTCGCAATGATTCGCTTCAACAACT 1448
Db 721 GAATATACCTCGTTGGTCCGGATTCCTCCAGTCAGCT 757

RESULT 14

AAx78003
ID AAX78003 standard; cDNA; 510 BP.

XX AAX78003;

DT 19-AUG-1999 (first entry)

XX A. thaliana GnTI-homologue cDNA fragment.

DE GnTI; beta-1,2-N-acetylglucosaminyltransferase; glycoprotein; plant;
KW deficient; defective; detection; transgenic plant; sugar residue;
KW medicine; antigenicity; deglycosylation; ds.

XX Arabidopsis thaliana.

OS DE19754622-A1.

PN 10-JUN-1999.

XX 09-DEC-1997; 97DE-01054622.

XX 09-DEC-1997; 97DE-01054622.

XX (VSCH/) VON SCHAEWEN A.

XX Von Schaewen A;

XX WPI; 1999-338905/29.

XX P-PSDB; AAY08890.

XX Nucleic acid encoding plant N-acetylglucosaminyl transferase I useful for
PT generating plants producing glycosylated proteins of low antigenicity.

XX Claim 5; Page 27; 37pp; German.

XX This invention describes novel plant N-acetylglucosaminyl transferase I
CC (GnTI) proteins and their encoding nucleic acids. The nucleic acid
CC sequences of the invention may be used for recombinant production of the
CC encoded proteins, which are then used to raise antibodies (Ab) for
CC identifying plants with deficient or defective production of the protein.
CC They may also be used to detect such plants by hybridization and to
CC isolate related sequences from other plants or to generate antisense or
CC sense constructs for reducing/deleting GnTI protein activity in plants.
CC These transgenic plants may be used to produce glycoproteins with
CC minimal, uniform and defined sugar residues. Such glycoproteins are
CC useful in medicine and research, e.g. human glucocerebrosidase for
CC treating Gaucher's disease. Plants which are defective or deficient in
CC production of the GnTI protein can be made to produce glycoproteins with
CC minimal, uniform and defined sugar residues, of low antigenicity. Use of
CC these plants eliminates the need for the difficult isolation and
CC deglycosylation of native proteins or preparation in defective animal
CC cells. This sequence represents cDNA which encodes an Arabidopsis
CC thaliana GnTI homologue

XX SQ Sequence 510 BP; 143 A; 104 C; 114 G; 149 T; 0 U; 0 Other;
Query Match 19.2%; Score 334; DB 2; Length 510;
Best local Similarity 78.4%; Pred. No. 8e-91; Indels 0; Gaps 0;
Matches 400; Conservative 0; Mismatches 110;
Qy 424 ATCGAAATGTACAGATGCCAGTGGCTGTAGTGTATGTTAGTTCGCAATCGGCTGTAT 483
Db 1 ATCGAAAGCTTGGATCCCGAGTGGCGCTGTAGTGTATGTTAGTTCGCTGCGTGCAGTCGACAG 60
Qy 484 TACCTGAAAAGACTATTAATCCATCTTAAATACCAATATCTGTTCGCTCAAAATAT 543
Db 61 TATCTTGAAGAGACTGTAAATCACTTAAATCAATCAAACTCCCGTTGCTTCAAAATAT 120
Qy 544 CCTCTTTTCAATCCAGGATGGATCACATCCTGATGTACAGAAAGCTTGTCTTTCAGCTAT 603
Db 121 CCTCTATTTATCTCAGGATGGATCTGATCAAGCTGTCAAGAGCAAGTCAATTCAGCTAT 180
Qy 604 GATCAGCTGACGTATATGCAGCACTTGGATTTTGAACCTGTGCATATCTGAAGACACAGG 663
Db 181 AATCAATTAACATATATGCAGCACTTGGATTTTGAACCACTGTCACTGAAGGCGCTGGC 240
Qy 664 GAGCTGATTCATACATAAAATTTGCAGCTCAATTACAAGTGGGCAATTCGATCAGCTGTTT 723
Db 241 GAAGTACTGCTACTACAGATTCGAGCTCAGCTACAGAGTGGGCACTGCAGCCAGTTGTTT 300
Qy 724 TACAAGCATATTTTAGCGGTGTATCATCTAGAAAGATGATATGAAATATGCCCTCAT 783
Db 301 TACAACACAAATTTAGTCGAGTGAATATATAGAGATGATATGGAATATGCTCCAGAC 360
Qy 784 TTTTTCGACTTTTTCGAGGCTGGAGCTACTCTTCTTGACAGAGCAAGTCGATTAATGCT 843
Db 361 TCTTTGATTACTTTGAGGCTGCAGCTAGTCTCATGGATAGGATAAAACCATTAATGCT 420
Qy 844 ATTCTCTTGGAAATGACAATGGACAAATGAGCTTTGTCCAAGATCCTTATGCTCTTTAC 903
Db 421 GCTTCATCATGAGACTGATAATGGACAGAGAGTTTGTGCATGATCCCTATGCGCTATAC 480
Qy 904 CGCTCAGATTTTTCGCGGTCTTGGATGG 933
Db 481 CGATCAGATTTCTTCCCTGCGCACGCTGG 510

RESULT 15

ADS70553
ID ADS70553 standard; cDNA; 278 BP.

XX AC

XX ADS70553;

XX 18-NOV-2004 (first entry)

XX Corn seedling-derived polynucleotide (cpds), SEQ ID 5569.

XX Corn; seedling-derived polynucleotide; cdp; ss; plant; seed growth;
KW seed development; disease resistance; insect infestation; fungal disease;
KW bacterial infection; Goss' Bacterial wilt; blight;
KW Stewart's bacterial wilt; Holcus spot; bacterial leaf blight; leaf spot;
KW bacterial stripe; maize dwarf mosaic virus infection;
KW environmental stress; water stress; pH stress; temperature stress;
KW pollution; injury; pesticide.

XX Zea mays.

XX US2003237110-A9.

XX 25-DEC-2003.

XX 06-AUG-2001; 2001US-00923876.

XX 12-MAY-1998; 98US-0085331P.

XX 21-APR-1999; 99US-00298329.

XX

(INCY-) INCYTE PHARM INC.
Lalgudi RV, Ito LY, Sherman BK;
WPI; 2002-195165/25.
New corn seedling-derived polynucleotides and polypeptides, useful in
identifying and altering desired characteristics associated with growth
and development, disease resistance, environmental adaptability, quality
and yield.
Claim 1; SEQ ID NO 5569; 33pp; English.
The invention relates to a corn seedling-derived polynucleotide (cdp)
selected from ADS64985-ADS71316, or their complements and fragments. Also
included are a composition for the detection of altered expression of a
cdp (comprising a polynucleotide selected from ADS64985-ADS71316), a
method of detecting a polynucleotide in a biological sample using a cdp,
a method for using oligomers (and amplification) to recover a regulatory
element from a DNA library using oligomers designed against a cdp, a
seedling specific regulatory element that regulates the expression of a
cdp, an expression vector containing a cdp or regulatory element, a plant
transformed with the vector, a host cell containing the vector (and
expressing a corn seedling derived protein, CDP), an anti-CDP antibody,
identifying a compound which binds a CDP and screening a plurality of
compounds for binding to cdp polynucleotide. The cdp polynucleotides,
proteins, vectors, cells and antibodies are useful for the
identification, evaluation and alteration of seed growth and development,
disease resistance (e.g. to insect infestation, fungal disease, bacterial
infection, Goss' Bacterial Wilt, blight, Stewart's bacterial wilt, Holcus
spot, bacterial leaf blight, leaf spot, bacterial stripe and maize dwarf
mosaic virus infection) and resistance to environmental stress (e.g.
water stress, pH stress, temperature stress, pollution, injury or
pesticides. The present sequence is cdp cDNA sequence.
Sequence 278 BP; 87 A; 52 C; 60 G; 79 T; 0 U; 0 Other;
Query Match 10.5%; Score 182; DB 7; Length 278;
Best Local Similarity 78.4%; Pred. No. 1.4e-44;
Matches 218; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
Qy 625 CACTTGGATTGTTGAACCTGTGCATCTGAAGACCGAGGAGCTGATTGCATCTACAAA 684
Db 1 CATGTGATCTTGACCTGTGCAAACTGAAGAGCGGAGAAATTGACGATATTACAAG 60
Qy 685 ATTGCAGCTCATTACAAGTGGGCATTTGGATCAGCTGTTTTACAAGCATAATTTAGCCGT 744
Db 61 ATTGCTAAACATTATTAAGTGGCCTTGGACAACCTATTTCATTAACATAATTTTGCTCGA 120
Qy 745 GTTATCATCATAGAAGATGATATGGAATTCGCCCTGATTTTTTTGACTTTTGGAGGCT 804
Db 121 GTAATCATCTTGAAGATGATGATGAGATTGCCCGGATTTTTTCGACTACTTCGAGGCT 180
Qy 805 GGAGCTACTCTCTTTCAGAGACAAGTCGATTATGGCTATTCTCTTTCGAATGACAAT 864
Db 181 GCAGCTAACTTCTTATGATATGATGACAGACGATATGGCTGTTCTGCTTGAATGACAAT 240
Qy 865 GGACAAATGCAAGTTTGTCAGAGATCCTTTATGCTCTTTA 902
Db 241 GGGCAGAAGCAGTTTGTGTTAACGACCCAAAGCTCTTTA 278
Search completed: December 14, 2005, 08:52:03
Job time : 1011 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 08:35:23 ; Search time 1307 Seconds
(without alignments)
10989.984 Million cell updates/sec

Title: US-10-615-144-3
Perfect score: 1737
Sequence: 1 gaattcgccgcgcattga.....ataatcgccgcgcgaattc 1737

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA Main:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1737	100.0	1737	8	US-10-615-144-3
2	1255	72.3	1669	8	US-10-615-144-1
3	767.2	44.2	1854	8	US-10-615-144-5
4	714.2	41.1	1836	7	US-10-425-114-29332
5	668.2	38.5	1594	7	US-10-437-963-21004
6	661.8	38.1	1882	7	US-10-425-114-487
7	661.8	38.1	2034	8	US-10-425-115-25820
8	661	38.1	1211	8	US-10-767-799-2499
9	614.6	35.4	2293	7	US-10-424-599-93080
10	581.2	33.5	1898	8	US-10-739-930-5511
11	577	33.2	1414	7	US-10-424-599-93082
12	527.8	30.4	954	7	US-10-424-599-93078
13	343.6	19.8	1849	7	US-10-437-963-21003
14	222	12.8	769	7	US-10-767-701-5996
15	182	10.5	278	3	US-09-923-876-5569
16	182	10.5	278	3	US-09-923-876-5569
17	181.8	10.5	2580	5	US-10-087-192-374
18	181.8	10.5	31477	5	US-10-087-192-373
19	146.4	8.4	2670	7	US-10-185-475-34
20	146.4	8.4	2670	9	US-10-713-970-12
21	146.4	8.4	2753	5	US-10-087-192-377
22	146.4	8.4	2937	8	US-10-844-874-13
23	146.4	8.4	2937	9	US-10-887-553A-329

24	146.4	8.4	4677	7	US-10-185-475-36	Sequence 36, Appl
25	146.4	8.4	4677	9	US-10-713-970-14	Sequence 14, Appl
26	146.4	8.4	32323	5	US-10-087-192-376	Sequence 376, Appl
27	136.6	7.9	2367	10	US-11-097-143-31280	Sequence 31280, A
28	122.6	7.1	590	7	US-10-021-323-15672	Sequence 15672, A
29	122.6	7.1	590	8	US-10-767-795-2500	Sequence 2500, Ap
30	105.8	6.1	509	7	US-10-767-701-21087	Sequence 21087, A
31	69.8	4.0	347	3	US-09-960-352-2775	Sequence 2775, Ap
32	58.8	3.4	4721	10	US-11-097-143-31279	Sequence 31279, A
33	58.6	3.4	613	8	US-10-363-345A-30605	Sequence 30605, A
34	58.6	3.4	613	8	US-10-363-345A-30606	Sequence 30606, A
35	58.6	3.4	613	9	US-10-363-483A-30605	Sequence 30605, A
36	58.6	3.4	613	9	US-10-363-483A-30606	Sequence 30606, A
37	55.2	3.2	587	8	US-10-357-930-45887	Sequence 45887, A
38	55	3.2	403	8	US-10-357-930-16061	Sequence 16061, A
39	39.6	2.3	613	8	US-10-363-345A-30607	Sequence 30607, A
40	39.6	2.3	613	8	US-10-363-345A-30608	Sequence 30608, A
41	39.6	2.3	613	9	US-10-363-483A-30607	Sequence 30607, A
42	39.6	2.3	613	9	US-10-363-483A-30608	Sequence 30608, A
43	38.8	2.2	535	5	US-10-027-632-91201	Sequence 91201, A
44	38.8	2.2	535	5	US-10-027-632-304430	Sequence 304430, A
45	38.8	2.2	535	6	US-10-027-632-91201	Sequence 91201, A

ALIGNMENTS

RESULT 1
US-10-615-144-3
; Sequence 3, Application US/10615144
; Publication No. US20040181827A1
; GENERAL INFORMATION:
; APPLICANT: Von Schaeven, Antje
; TITLE OF INVENTION: Plant GntI Sequences and the Use Thereof for the Production
; TITLE OF INVENTION: of Plants Having Reduced or Lacking N-acetyl Glucosaminyl
; TITLE OF INVENTION: Transferase I(GntI) Activity
; FILE REFERENCE: 032266-003
; CURRENT APPLICATION NUMBER: US/10/615,144
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/591,466
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: EP 98/08001
; PRIOR FILING DATE: 1998-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Nicotiana tabacum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (733)...(741)
; OTHER INFORMATION: function: Asn codon in this context is a potential
; OTHER INFORMATION: glycosylation site;
; OTHER INFORMATION: product: N-glycosylation consensus sequence;
; OTHER INFORMATION: phenotype: N-glycans modulate protein properties;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (733)...(741)
; OTHER INFORMATION: standard name: N-glycosylation site;
; OTHER INFORMATION: label: pot-CHO;
; OTHER INFORMATION: note: GntI sequences from animals do not contain
; OTHER INFORMATION: this feature.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (127)...(1467)
; OTHER INFORMATION: codon start: 127;
; OTHER INFORMATION: function: initiates complex N-glycans on secretory
; OTHER INFORMATION: glycoproteins;
; OTHER INFORMATION: EC_number: 2.4.1.101;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (127)...(1467)

= instant gghh

Db 1561 ACAAGCTTTTAAGTATACGAGGAGCTCACTGTTCTAGTGTGAAGGATATCGGCTT 1620
Qy 1621 CTTAGTATTGATGAATCATCAACACACCTATTATTTTAAAGTGTTCAGAACATAAGAG 1680
Db 1621 CTTAGTATTGATGAATCATCAACACACCTATTATTTTAAAGTGTTCAGAACATAAGAG 1680
Qy 1681 GAAATGTAGCCCTGTAAAGACTATATACATGGGACCATCATATCGCGGCGGGAATTC 1737
Db 1681 GAAATGTAGCCCTGTAAAGACTATATACATGGGACCATCATATCGCGGCGGGAATTC 1737

RESULT 2

US-10-615-144-1
; Sequence 1, Application US/10615144
; Publication No. US2004018127A1
; GENERAL INFORMATION:
; APPLICANT: Von Schaeuwen, Antje
; TITLE OF INVENTION: Plant GntI Sequences and the Use Thereof for the Production
; TITLE OF INVENTION: of Plants Having Reduced or Lacking N-acetyl Glucosaminyl
; TITLE OF INVENTION: Transferase I(GntI) Activity
; FILE REFERENCE: 032266-003
; CURRENT APPLICATION NUMBER: US/10/615,144
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/591,466
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: EP 98/08001
; PRIOR FILING DATE: 1998-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1669
; TYPE: DNA
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (659)...(667)
; OTHER INFORMATION: function: Asn codon in this context is a potential
; OTHER INFORMATION: glycosylation site;
; OTHER INFORMATION: product: N-glycosylation consensus sequence;
; OTHER INFORMATION: phenotype: N-glycans modulate protein properties;
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (659)...(667)
; OTHER INFORMATION: standard_name: N-glycosylation site;
; OTHER INFORMATION: label: pot-CHO;
; OTHER INFORMATION: note: GntI-coding sequences from animals do not
; OTHER INFORMATION: contain this feature.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53)...(1393)
; OTHER INFORMATION: codon_start: 53;
; OTHER INFORMATION: function: initiates complex N-glycans on secretory
; OTHER INFORMATION: glycoproteins;
; OTHER INFORMATION: EC_number: 2.4.1.101;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53)...(1393)
; OTHER INFORMATION: product: beta-1,2-N-acetylglucosaminyltransferase I;
; OTHER INFORMATION: evidence: EXPERIMENTAL;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (53)...(1393)
; OTHER INFORMATION: gene: cgl;
; OTHER INFORMATION: standard_name: gntI;
; OTHER INFORMATION: label: ORF;
; OTHER INFORMATION: note: first gntI sequence from potato (unpublished).
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (15)...(52)
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (1394)...(1655)
; FEATURE:

; NAME/KEY: CDS
; LOCATION: (80)...(139)
; OTHER INFORMATION: function: membrane anchor (amino acids 10-29);
; OTHER INFORMATION: product: hydrophobic amino acid stretch in GntI;
; OTHER INFORMATION: standard_name: membrane anchor of a type II Golgi
; OTHER INFORMATION: protein;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (80)...(139)
; OTHER INFORMATION: note: identified by comparison with GntI sequences
; OTHER INFORMATION: from animals.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(14)
; OTHER INFORMATION: function: used for cloning the cDNA library in
; OTHER INFORMATION: Lambda ZAPII;
; OTHER INFORMATION: product: EcoRI/NotI-cDNA adapter;
; OTHER INFORMATION: number: 1.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1656)...(1669)
; OTHER INFORMATION: product: EcoRI/NotI-cDNA adapter;
; OTHER INFORMATION: number: 2.
; US-10-615-144-1
Query Match 72.3%; Score 1255; DB 8; Length 1669;
Best Local Similarity 89.1%; Pred. No. 0;
Matches 1412; Conservative 0; Mismatches 140; Indels 33; Gaps 4;
Qy 126 GATGAGGGGAACAAGTTTCTGCTGATTCCTCGGTACTCTCTCATCTTGGCTGCTGCTGC 185
Db 52 GATGAGGGGAACAAGTTTCTGCTGATTCCTCGGTACTCTCTCGTGGCTGCTCTCGC 111
Qy 186 CTTTCATCTACACAGATCGCGCTTTTTCGCACACAGTCAGAAATATGAGATCGCTTGC 245
Db 112 CTTTCATCTACACAGATCGCGCTTTTTCGCACACAGTCAGAAATATGAGATCGCTTGC 171
Qy 246 TGCTGCAATTGAAGCAGAAAATCATTTGTACAGCCAGACCCAGATTGCTTATGACCAAT 305
Db 172 TGCTGCAATTGAAGCAGAAAATCATTTGTACAGTCAGACCAAGATTGCTTATGACCAAT 231
Qy 306 TAGCCTGCAGCAAGAAAGATAGTTGCTCTTGAAGAACAAATGAAGCGTCAGGACAGGA 365
Db 232 TAGCCAGCAGCAAGGAGAGTAGTGTCTTGAAGAACAAATGAAGCATCAGGACAGGA 291
Qy 366 GTGCCGCAATTAAGGGCTCTTGTTCAGGATCTTGAAGTAGGGGCATATAAAGTTGAT 425
Db 292 GTGCCGCAATTAAGGGCTCTTGTTCAGGATCTTGAAGTAGGGGCATATAAAGTTGAT 351
Qy 426 CGGAAATGTACAGATGCCAGTGGCTGTGTAGTTGTTATGGCTTGCATTCGGGCTGATTA 485
Db 352 CGGAAATGTACAGATGCCAGTGGCTGTGTAGTTGTTATGGCTTGCATTCGGGCTGATTA 411
Qy 486 CTTGGAAGAACTATTAAATCAATCTTAAATACCAATATCTGTGGTCAAAATATCC 545
Db 412 CTTGGAAGAACTATTAAATCAATCTTAAATACCAATATCTGTGGTCAAAATATCC 471
Qy 546 TCTTTTCATATCCAGGATGGATCATCTCATCTGAGGAGCTTGTTCAGGCTATCA 605
Db 472 TCTTTTCATATCCAGGATGGATCAATCTTAAATACCAATATCTGTGGTCAAAATATCC 531
Qy 606 TCAGCTGAGCTATATGACAGCTTGGATTTTGAACCTGTGCTATCTGAAAGACCGAGGA 665
Db 532 TCAGCTGAGCTATATGACAGCTTGGATTTTGAACCTGTGCTATCTGAAAGACCGAGGA 591
Qy 666 GCTGATTGCTACTACAAAATTCACGTCATTAAGTGGGCAATGGATCAGCTGTTTGA 725
Db 592 ACTGGTTGCATACTACAAAGATTGCACGTCATTACAAAGTGGGCAATGGATCAGCTGTTTGA 651
Qy 726 CAAGCATATTTTAGCCGTGTTTATCATCTACTAGAAAGATGATGGAATTCCTGATTT 785
Db 652 CAAGCATATTTTAGCCGTGTTTATCATCTACTAGAAAGATGATGGAATTCCTGATTT 711

Best Local Similarity 73.1%; Pred. No. 2.8e-216; Mismatches 363; Indels 0; Gaps 0;									
Qy	139	AGGTTTCTGCTGATCTCCGGTACCTCCTCATCTTGGCTGCTGCGCTTCATCTACACA	198						
Db	141	AGGATCTCGTGTGACTTGGAGATTTCTTCTCATCCGGCAGCTTTTCATGTTTCATCTACATC	200						
Qy	199	CAGATGGCGCTTTTGGACACACAGTCAAGATATCAGATCGCTTGTCTGCTGCAATTTGAA	258						
Db	201	CAGATGAGGCTTTTCCAGACCAATCACAGTATGCAGATCGCTTCAGTTCCGCTATCGAA	260						
Qy	259	GCAGAAATCATTTGTAACAAGCCAGACCCAGATTTGCTTATTTGACCAGATTTAGCCTCGACAA	318						
Db	261	TCTGAGAACCAATGCACTAGTCAATCGAGGCTCATAGATGAATTTAGCATCAACACAG	320						
Qy	319	GGAAAGATAGTGTCTTTGAAGAACAAATGAAGCGTCAAGACCAAGAGTGCAGCAATTA	378						
Db	321	TCGCGGATTTGTCCTCGAAGATATGAAGAACCCGACGACGAAGAACTTTGTGCAGCTT	380						
Qy	379	AGGCTCTTGTTCAGGATCTTGAAGTAAGGCAATAAAGATTTGATCGGAAATGTACAG	438						
Db	381	AAGGATCTAATCCAGACGCTTTGAAGAAAAAGAAATAGCAAACTCCTCAAGGTGGACAG	440						
Qy	439	ATGCCAGTGGCTGTGTAGTTGTTATGGCTTGCATCGGCTGATTTACTTGAAAAAGACT	498						
Db	441	ATGCTGTGGCTGTGTAGTGTGTTATGCGCTGCAGTCTGTGCAGACTATCTTTGAAGGACT	500						
Qy	499	ATTAAATCCATCTTAAATAACAAATATCTGTGTGGTCAAAATATCTCTTTTCATATCC	558						
Db	501	GTAAATCAGTTTAAACATATCAAACTCCGTTGCTTCAAAATATCTCTATTTATATCT	560						
Qy	559	CAGATGATCACATCTCTGATGTCAAGAGCTTGTCTTTGAGCTATGATCAGTCAAGAT	618						
Db	561	CAGATGATCTGATCAAGCTGTCAAGAGCAAGTCAITGAGCTATTAATCAATTAACATAT	620						
Qy	619	ATGCAGCACTTGGATTTTGAACCTGTGCATATCTGAAGACACAGGAGGCTGATTCATAC	678						
Db	621	ATGCAGCACTTGGATTTTGAACCTGTGTCAAGAGGCTGTGTGAGTCACTGCGTAC	680						
Qy	679	TACAAATTTGACGTCATTTACAAGTGGGATTTGGATCAAGTGTGTTTACAAGCAATTTT	738						
Db	681	TACAAATTTGACGTCATTTACAAGTGGGATTTGGATCAAGTGTGTTTACAAGCAATTTT	740						
Qy	739	AGCGTGTATCATACTAGAGATCATATGNAATTTGCCCTGATTTTTCAGCTTTT	798						
Db	741	AGTCAGTGTATCATACTAGAGATCATATGNAATTTGCCCTGATTTTTCAGCTTTT	800						
Qy	799	GAGGCTGGAGTACTCTTCTTACAGACACAGTGTGATTTGGCTATTTCTTTGGAAT	858						
Db	801	GAGGCTGGAGTACTCTTCTTACAGACACAGTGTGATTTGGCTATTTCTTTGGAAT	860						
Qy	859	GACAAATGGAATAGTGTCTTCAAGATCTTCAAGATCTTCAAGATCTTCAAGATGCGCA	918						
Db	861	GATAATGGAATAGTGTCTTCAAGATCTTCAAGATCTTCAAGATCTTCAAGATGCGCA	920						
Qy	919	CCCGTCTTGGATGATGCTTTCAAAATCTTCTTGGAGCAATTTATCTTCAAGTGGCA	978						
Db	921	CCTGGCTTGGTGGATGCTTCAAGATCTTCAAGATCTTCAAGATCTTCAAGATGCGCA	980						
Qy	979	AGGCTTACTGGAGCACTGGCTTGAAGTCAAGAGATTCAGAGGTCGACAAATTTAT	1038						
Db	981	AAAGGCTTACTGGAGCACTGGCTTGAAGTCAAGAGATTCAGAGGTCGACAAATTTAT	1040						
Qy	1039	CGCCCAAGATTTGACAGACATATAATTTTGGTGGAGCATGGTTCTAGTTTGGGGAGTTT	1098						
Db	1041	GCACCGAGTCTGTAGACATACAAATTTTGGTGAACATGGCTTAGTTTGGGACAGTTT	1100						
Qy	1099	TTCAAGCAGTATCTTGGAGCAATTAATTAATGATGTCCAGTTGATTTGGAAGTCAATG	1158						
Db	1101	TTCAAGCAGTATCTTGGAGCAATTAATTAATGATGTCCAGTTGATTTGGAAGTCAATG	1160						
Qy	1159	GACCTTAGTTACTTTTGGAGGACCAATTAAGTGAACACTTTTGGTGAATTTGTTATGAG	1218						

Db	1161	GACCTGGGATCTCTGACAGAGGAAACTATACCAAGTACTTTTCTGGCTTAGTGAGCAA	1220						
Qy	1219	GCTAAGCCCATCTCATGAGCTGATGCTGCTTGAAGAGCAATTTAACATAGATGATGTG	1278						
Db	1221	GCAGGACCAATTCAGGTTCTGACCTTGTCTTAAGGCTCAAAACATATAAGGATGATGAT	1280						
Qy	1279	CGTATTCAGTACAGAGATCAATAGACTTTTGAATAATTCGACGCGCAATTTGGCAATTTT	1338						
Db	1281	CGTATCCGGTATAAAGACCAAGTAGAGTTTGAACGCAATTCAGGCGAATTTGGTATATTT	1340						
Qy	1339	GAAGATGGAAGGATGCTGTACACGTCGACGATATAAAGCAATAGTAGTTTCCGGTAC	1398						
Db	1341	GAAGATGGAAGGATGCTGTGCGCACGACGATATAAAGCAATAGTAGTTTCCGAATC	1400						
Qy	1399	CAAACTCCAGAGCTGTATTCCTTTGTTGGCCATGATTCGCTTCAACCACTCGGAATGAA	1458						
Db	1401	CAGCAACAAGACGTTATTCCTGTTGGCCAGATTTCTGTATGCGAGCTTGGAAATCGA	1460						
Qy	1459	GATACCTTAACAAAGATATGATTCGAGGA	1486						
Db	1461	AATTCCTGATGCAAAACATATGAAAGGA	1488						
RESULT 4									
US-10-425-114-29332									
; Sequence 29332, Application US/10425114									
; Publication No. US2004003488A1									
; GENERAL INFORMATION:									
; APPLICANT: Liu, Jingdong									
; APPLICANT: Zhou, Yihua									
; APPLICANT: Kovalic, David K.									
; APPLICANT: Screen, Steven E									
; APPLICANT: Tabaska, Jack E									
; APPLICANT: Cao, Yongwei									
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With									
; FILE OF INVENTION: Plants and Uses thereof for Plant Improvement									
; FILE REFERENCE: 38-21(53313)B									
; CURRENT APPLICATION NUMBER: US/10425,114									
; CURRENT FILING DATE: 2003-04-28									
; NUMBER OF SEQ ID NOS: 73128									
; SEQ ID NO 29332									
; LENGTH: 1836									
; TYPE: DNA									
; ORGANISM: Glycine max									
; FEATURE:									
; OTHER INFORMATION: Clone ID: UC-GMFLMINSOV005B11_FLI									
US-10-425-114-29332									
Query Match 41.1%; Score 714.2; DB 7; Length 1836;									
Best Local Similarity 71.8%; Pred. No. 1.6e-200;									
Matches 1037; Conservative 0; Mismatches 288; Indels 120; Gaps 3;									
Qy	139	AGTTTCTGCTGATTTCCGGTACCTCCTCATCTTGGCTGCTGCTGCTTCACTTACACA	198						
Db	43	AAAGTTTCTGCTGATTTTCGGTTTCTTCTTGTATGTCAGCTGGCGTCTTTCATCTATATC	102						
Qy	199	CAGATGCGCTTTTGGACACACAGTCAGAAATATGAGATCGCTTGTGCTGCTCAATTTGAA	258						
Db	103	CAGATGCGCTTTTGGACACACATCAGAAATATGCTGATCGCTGCTGCTGCTGCTATCGAA	162						
Qy	259	GCAGAAATCATTTGTACAAAGCCAGACCCAGATTTGCTTATTTGACCATTTAGCCTCGACAA	318						
Db	163	GCTGAATACCATTTGTACAAAGTCAGATCGCTGATTTGATCAGATTTAGCTTGAACAA	222						
Qy	319	GAAGATAGTGTCTTGTGAAGAACAAAT	347						
Db	223	GCAGCAATTTGGCCCTCGAAGATTTAAATTTGCTGCTTGTGTAATTTCTTTCAGAAAGC	282						
Qy	348	GAAGCTCAGACCCAGAGGTCGACAAATTAAGGCTCTTCTTTCAGGATCTTGAAGATA	406						
Db	283	AGAAACGTCGAGACCAAGAAATGTCAGCAAAACAAAGTCTTCTGTAAGATCTTGAAGAA	342						
Qy	407	AGGGCAATAAAAGTTGATCGGAAATGTACAGATCGAGTGGCTGCTGTGTGTATGG	466						

QY 687 TGCAGCTCATTTACAGTGGGCTATGGATCAGCTGTTTACAGCATATTTTACCGGT 746
DB 555 AGCCAAACATTTACAAATGGGCTTGGATGAGATTTTATCAACACACAACTTTGGCGAGT 614
QY 747 TATCATATCTAGAAATGATATGAAATTTGCCCTGATTTTGTGACTTTTGTGAGGCTGG 806
DB 615 AATAATTTCTGGAGATGATATGGAGATCGCCCTGATTTTGTGACTACTTTGAGGCTGC 674
QY 807 AGCTACTCTTTGTACAGAGCAAGTCGATTTATGCTATTTCTTTTGGATGCAATGG 866
DB 675 AGCTAAATTTACTTGTATAACGATAAGCAATCATGCGCGTTTCTTCTGGAATGCAATGG 734
QY 867 ACATATGAGTTTGTCCAGATCTTTATGCTCTTTTACCGCTCAGATTTTTCCTCGGTCT 926
DB 735 ACAAAAGCAATTTGTATTAGACCCAAAGCTCTTTTACCGCTCGATTTTCTTCTGACT 794
QY 927 TGGATGATGCTTTTCAAAATCTTCAATTTGGGACGAATTTATCTCAAGTGGCCAAAGCTTA 986
DB 795 TGGATGGATTTTAAACAGCTTACATGGATTGAGCTGTACCTAAGTGGCCCAAGCTTA 854
QY 987 CTGGGACGACTGGCTTAAGACTCAAAAGAAATCACAGAGTGCACAAATTTATTCGCCCGA 1046
DB 855 TTGGGATGACTGGGTGAGCTAAAGAGAGGTACACAGAGACCGGCAATTTTATTCGCCCGA 914
QY 1047 AGTTTGCAGACATATATTTTGGTGGAGCTGTTCTAGTTTGGGCGAGTTTTCAGCA 1106
DB 915 AGTCTGCAGAACATCAATTTTGGGCAACATGGGTCAAGCATGGGCGCAATTTCTTCAAGCA 974
QY 1107 GTATCTTGGAGCAATTTAAACTAAATGATGTCAGGTTGATTTGGAAGTCAATGACCTTAG 1166
DB 975 GTACTTGGNACCAATCAAGCTTAATGATGCCCATATCAAGTGAATTTCTGAGGACCTGAG 1034
QY 1167 TTACCTTTTGGAGGCAATTAAGTGAACACTTTTGGTGTGCTTTGTTTAAAGGCTAAGCC 1226
DB 1035 CTACCTCAAGGAGGCAAGTCTCTGATCCAAATTTGGGAAGACGTCGCTAGTGGCACCC 1094
QY 1227 CATCATGAGCTGATGCTGCTTGAAGCAATTTAAGCATGATGCTGCTGATTTCA 1286
DB 1095 TCTCCATGATCTGATGCGCGTGTGAAGCCACATATGGAATGGAATGGAATGGAATG 1154
QY 1287 GTACAGAGTCAACTAGACTTTGAAATATCGCACGGCAATTTGGCAATTTTGAAGAATG 1346
DB 1155 GTATACGATCAGGAGACTTTGAGCGGATAGCTGCTCAATTTGGAATTTTGAAGAATG 1214
QY 1347 GAAGGATGCTGACCACTGACGATATTAAGGAATAGTATTTTCCCGTACCA---AAC 1403
DB 1215 GAAGGATGCAATTCACAAAGACGCTTATAAGGAGTAGTATCTTCCGCTACAAGAGTAG 1274
QY 1404 GTCCAGAGCTGATTTCTTGTGGCCATGATTCGCTTCAACACT 1448
DB 1275 TCGAAGACGAATATACCTGTTGGTCCGATTTCCCTCAGTCAGCT 1319

RESULT 6

US-10-425-114-487
; Sequence 487, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 487
; LENGTH: 1882

; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700077454_F11
US-10-425-114-487

Query Match 38.1%; Score 661.8; DB 7; Length 1882;
Best Local Similarity 68.4%; Pred. No. 5.9e-185;
Matches 931; Conservative 0; Mismatches 427; Indels 3; Gaps 1;

QY 147 CTGTGATTTCCGGTACCTCTCATCTTGGCTGCTGCGCTTCATCTACACACAGATCGG 206
DB 160 CTGCGACTCCGCT 219
QY 207 GCTTTTTCGACACAGCTCAGAAATATGCAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 266
DB 220 CCTCTTGTCTACTAGTCCCATATGCGGACCGTCTTGGGAAGCAGAAAGATCTGAAA 279
QY 267 TCATTTGTAACAGCCAGACAGATTTGCTTATTTGACCAAGATTAGCTTGCAGCAAGGAGAT 326
DB 280 TCAATGCACAGTCAATTTAAAGTCTTGTATCGATCAAGTCAGCATGCGACGAGGAAGAT 339
QY 327 AGTTGCTCTTGAAGCAAAATGAAGCTCAGCAACGAGGTCGCGACAAATTTAAGGGCTCT 386
DB 340 TGTAGCATTTGGAAGAGATTAAGGTACGGAAGATGAAGAACGTGCACATCTGAGGATTTT 399
QY 387 TGTTCAGGATCTTGAAGTAAAGGCGATATAAAGTTGATCGGAAATGTACAGATGCCAGT 446
DB 400 GATTAAGGATCTTGAANAAGGAGTGTGCAGAGCTACTAGACAAGATGTGTTCCCGT 459
QY 447 GGCTGCTGATGTTGTTATGGCTTGCATTCGGGCTGATTAACCTGGAAGAAAGATTTAAATC 506
DB 460 TGCTGCTGTTGTCATAATGGCTTGCATTCGACCCAGACTATTTAGAGAGAACAGATTGAATC 519
QY 507 CATCTTAAATACCAAAATCTGTTGCTGCTCAAAATATCTCTTTTTCATATCCACGATGG 566
DB 520 TATCTGAAGTATCAGACATCAGTGTCTTCCGAGTTTCCGCTTTTATCTCAGATGG 579
QY 567 ATCATCTCTGATGTGAGAACTTGTCTTGTAGCTATGATCAGCTGAGCTATATATCAGCA 626
DB 580 AGCAATCGAGCTGTGAAGATTAAGCTTTGGAATATAAACAATAACATACATCAGCA 639
QY 627 CTGGAATTTGAACTGTGCATATCTGAAGACCCAGGGAGCTGATGATGATATCAAAAT 686
DB 640 TGTGATCTTGAACCTGTGCAAACTGAAGGCGCAGAGAAATGACAGCATATTTACAAGAT 699
QY 687 TGCAGCTCATTTACAAGTGGCATTTGGATCAGCTGTTTACAAGCATATTTTACCGCTGT 746
DB 700 TGCTAAACATTTATAGTGGGCTTGGCAACCTATTTCAATTAACATAATTTTGTCTCGAGT 759
QY 747 TATCATATCTAGAAGATGATATGAAATTTGCCCTGATTTTGTAGCTTTTGTAGGCTGG 806
DB 760 AATCATTTCTAGAAGATGACATGGAGATTTGCCCGGATTTTTCGACTACTTTCGAGGCTGC 819
QY 807 AGCTACTCTTCTTGAAGAGCAAGTCAATTAATGCTATTTCTTCTTGAATGCAATGG 866
DB 820 AGCTAAACTACTTGAATATGACAGCATATGCTGTTTCTGCTTCTGGAATGCAATGG 879
QY 867 ACAATGCAATTTGTTCGAAGATCTTATGCTCTTTACCGCTCAGATTTTTCCTCGGCT 926
DB 880 GCAGGAAGCAGTTTGTAAACGCCCAAAAGCTCTTTTACCGCTCAGATTTTCTTCTCGGCT 939
QY *927 TGGATGAGTCTTCAAAATCTACTTTGGGACCAATTTATCTCAAAAGTGGCCAAAGCTTA 986
DB 940 TGGATGAGTGTAAACAAAGTCCATTTGGAATTTGTCACCAAAAGTGGCTTAAGCTTA 999
QY 987 CTGGGACGACTGGCTTAAGACTCAAAAGAGAAATCAGAGGTGCAGCAATTTTATTCGCCAGA 1046
DB 1000 TTGGGATGATTTGGGTGAGGCTTAAGAGAGGTACATGAGCATCGCAATTTATACGCCGGA 1059
QY 1047 AGTTTGCAGAAATATAATTTTGGTGGAGCATGGTCTTAGTTTGGGCGAGCTTTTTCAGCA 1106
DB 1060 AATTTGTAGAACATACAAATTTTGGCAAGCATGGATCAAGCTTTGGGACAGCTTCTTCAGCA 1119

Qy	1107	GTATCTTGAGCCCAATTAAACTRAAATGATGTCAGGTTGATTGGAAGTCAATGGACCTTAG	1166
Db	1120	ATACTTGGAGCCCAATTAAAGTTAAACGATGTTCAAAATCGATTGGAATTTCTGAGGATCTGAG	1179
Qy	1167	TTACCTTTTGGAGGACAAATTACGTGAAACACATTTGGTGAATTTGTTAAAAAGGCTAAGCC	1226
Db	1180	CTACCTTTAGGGAGGATAAGTTTTCACCCAAATTTTGGAAAGAGTGGCTAGTGCACCTCC	1239
Qy	1227	CATCCATGGAGCTGATGCTGCTTTGAAAGCAATTTAACTATAGATGGTATGTCGATTTCA	1286
Db	1240	TCTCCGTGATCAGATGCTGTTTGAAGCCCAATATGGCTGAGGATGTAAGATCCA	1299
Qy	1287	GTACAGAGATCAACTAGACTTTTGAAATATCGCAGGCAATTTGGCATTTTGTGAAGAATG	1346
Db	1300	ATACGACGACCAAGAAGTTTCGAGCAGATAGCTCGTCAATTTGGAAATTTTCGAAGATG	1359
Qy	1347	GAAGGATGGTGTACCAAGCTGACGATATTAAGAGATATAGTAGTTTCCGGTACCA---	1403
Db	1360	GAAGGATGGCATCCCAAGAACCGCTTTACAAGAGAGTGGTGTCTTCCGGTACAACAGTAG	1419
Qy	1404	GTCAGACGTGATTTCTCTTTGGCCATGATTCGCTTCAACAACTCGCAATTTGAAGATAC	1463
Db	1420	TCAAAGCGGAAATTTCTGTCAGCCCAAGATTTCTTCTGCTAGCTGGGGGTGATGCAATG	1479
Qy	1464	TTAACAAAGATATGATTGAGGAGCCCGGCAAAATTTTGG	1504
Db	1480	TGAGAAGGTAGATTTTACCTTGGAAATGGCAAAATATCATG	1520
RESULT 7			
US-10-425-115-25820			
; Sequence 25820, Application US/10425115			
; Publication No. US20040214272A1			
; GENERAL INFORMATION:			
; APPLICANT: La Rosa, Thomas J.			
; APPLICANT: Kovalic, David K.			
; APPLICANT: Zhou, Yihua			
; APPLICANT: Cao, Yongwei			
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with			
; TITLE OF INVENTION: Plants			
; FILE REFERENCE: 38-21(53222)B			
; CURRENT APPLICATION NUMBER: US/10/425,115			
; CURRENT FILING DATE: 2003-04-28			
; NUMBER OF SEQ ID NOS: 369326			
; SEQ ID NO 25820			
; LENGTH: 2034			
; TYPE: DNA			
; ORGANISM: Zea mays			
; FEATURE:			
; OTHER INFORMATION: Clone ID: MRT4577_12355C.1			
US-10-425-115-25820			
Query Match 38.1%; Score 661.8; DB 8; Length 2034;			
Best Local Similarity 68.4%; Pred. No. 6.2e-185;			
Matches 931; Conservative 0; Mismatches 427; Indels 3; Gaps 1;			
Qy	147	CTGTGATTTCCGGTACCTCCTCATCTTGGCTGCTGCTGCTTCATCTACACAGATGCG	206
Db	205	CTCGACCTCCGGCTCCTCCTCCTGCTGAGCTGCGATTCATCTACATCCCAAGTGGC	264
Qy	207	GCTTTTGGACACAGTCAGAAATATGACATCGCTTGTGCTGCTGCAATTTGAAGCAGAAA	266
Db	265	CCTCTTTGCTACTCAGTCCCATATGCGGACCGCTTTGCGGAAGCAGAAAGATCTGAAA	324
Qy	267	TCATTTGACAAAGCCAGATTTGCTTTATTGACCAAGATTTAGCTTCGACGAAAGAAAT	326
Db	325	TCAATGCACCAAGTCAATTAAGTCCCTTGATCGATCAAGTCAGATGACGACGAGGAAGAT	384
Qy	327	AGTTGCTCTTGAAGACAAATGAAGCGTCAGACCAAGAGTGGCCGACAAATTAAGGGCTCT	386
Db	385	TGTAGCATTTGGAAGAGATAAAGGTACGGCAAGATGAAGAACGTGCACATCTGAGGATTTT	444

```
Db 1525 TGAGNAGTAGATTTTACCCCTTGGAATGGCAATATCATG 1565
|||||
RESULT 8
US-10-767-795-2499
; Sequence 2499, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53234)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 2499
; LENGTH: 1211
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C89123_1
US-10-767-795-2499

Query Match 38.1%; Score 661; DB 8; Length 1211;
Best Local Similarity 75.1%; Pred. No. 7.8e-185;
Matches 851; Conservative 0; Mismatches 280; Indels 2; Gaps 2;

QY 127 ATGAGAGGGAACAAGTTTGTGTCATTTCCGGTACCTCCTCATCTTGCTGCTCGCC 186
DB 80 ATAAAATGGCTAAGCTAAGCTGCGATTTCCGGTACCTCCTCATCTCGCTGCTGTA 139
|||||
QY 187 TTCATCTACACAGATGCGGCTTTTGGACACAGTCAGATATGCAATATGCAATCGCTTGT 246
DB 140 TTCATATACATCCAGATGCGGCTTTTCCACAGCAATCTGAATATGCAATCGAATGGCA 199
QY 247 GCTGCATTTGAAGCAAAATCATTTGTACACGACAGCAGATTCCTTATTCACAGATT 306
DB 200 GAAGCGGTGATGACAGACCAATTTGTACGAGTCAATTCGACTACTAATTTGATCAAAAT 259
QY 307 AGCTGCGACAGGAGGATAGTGTCTTTGAAGCAAAATCAAGCGTCAGACACAGGAG 366
DB 260 AGTATGCAACAGAGCAAAATCGTGGCCCTTGAAGAGCGGGAAGAGCAAGGATCAGGAG 319
QY 367 TGCCGACAAATTAAGGCTCTTGTTCAGGATCTTTGAAGTAAAGGCAATAAAAGTTGATC 426
DB 320 TGGCACAATTTGAAGACTCTTGTATATGATCTTTGAAGAAAGCTCTTCAAGAGATAAT 379
QY 427 GGAATGTACAGATGCGAGTGTGCTGTAGTTGTATGCTTGCATCGGCTGATTTAC 486
DB 380 GACAAACACAGATACCTGTGGCAGCTGTAGTAATCATGCGGTGCAATCGTCTGATTAT 439
QY 487 CTGGAAGACTATTAAATCCATTTAAATACCAATATCTGTGCGTCAAAATATCTCT 546
DB 440 CTGGAAGGACAGTTCGCTGTCTTAAAGTATCAGAGCTCTGTGCTTGAAGTATCA 499
QY 547 CTTTTCATATCCAGATGGATCACAATCTCTGTATGTGAGGAAGCTGTCTTTGAGCTATGAT 606
DB 500 CTTTGTGTATCTCAGATGATGATCAATCCAGGGTTAAAGCTTTGAGTTATAG 559
QY 607 CAGTCACTATATGAGCACTTGGATTTTGAACCTGTGCTATCTGAAAGACCCAGGGAG 666
DB 560 GAGCTAACTATATGAGCAGACATAGATATGATCCCGTGCATACAGACCGGCTGGGAA 619
QY 667 CTGATTCATCTACAAAATTCACCTCATTAAGTGGGCAATGGAATCAGCTGTTTTAC 726
DB 620 TTGATCGCATCTACAAAGATTTGCCGCTCACTACAAATGGGCATTTGATGATTTCTAC 679
QY 727 AAGCAAAATTTAGCCGCTGTTATCATCTAGAGATGATATGGAATTTGCCCTGATTTT 786
DB 680 AAGCAAAATTTAGCCGAGTAAATTAATCTTGAAGATGATATGGAATTTGCCCTGATTTT 739
```

```
QY 787 TTTGACTTTTTTGGGCTGGAGCTACTCTTTTGACAGAGACAGTTCGATTTATCGCTATT 846
DB 740 TTTGATTACTTTGGGCGAGCTGCTGCCCTTCTCGACAAAGACAGTCAAATTTATGGCTTT 799
QY 847 TCTTCTTTGGAATGACATGACAAAATGCAAGTTTGTCCAAGATCTTTATGCTTTTACCGC 906
DB 800 TCCTCATGGAATGACAAATGGGCAAAAGCAGTTTGTGTATGACCCCATATGCACTTTATCGC 859
QY 907 TCAGATTTTTTCCCGGTCTTGGATGGATGCTTTCAAAATCTACTTTGGGACGAATTTATCT 966
DB 860 TCAGATTTTCTTCTGCTTGGCTGGATGCTTAAATCTGTATGGAATGAGCTATCA 919
QY 967 CCAAAGTGGCCAAAGGCTTACTTGGGACGACTGGCTAAAGCTCAAGAGAAATCAAGAGGT 1026
DB 920 CCAAATGGCCAAAGCTTACTTGGGATGACTGGTTGAGATTTAAAAAATAATCAACATGGT 979
QY 1027 CGACAATTTATTCGCCCAGAAAGTTTGACAGACATATATTTTGGTGAGCAGTGTCTAGT 1086
DB 980 CGACAATTTCTTCTGCTCTGGAATATGACAGACATATATTTTGGTGAGCATGGTTCAAGC 1039
QY 1087 TTGGGCGAGTTTTTCAAGCAGTATCTTTGAGCCAATTTAAACTTAAATGATGTCCAGTTGAT 1146
DB 1040 ATGGGCGAGTTTTTTCGAAAATACTTTGCACCTATTAAAGATGAATGAGCT-CAGGTGAC 1098
QY 1147 TGAAGTCAATGAGACCTTTAGTTACTTTTGGAGACAAATTAAGTGAACACACTTTGGTGAC 1206
DB 1099 TGAAGTCCA-GGATTTGAGCTACTTAACCTGAGGAAAAATATGCCCAATACTTTGACAGC 1157
QY 1207 TTGGTTAAAAAGCTAAGCCCATCATGGAGCTGATGCTGTCTTTGAAAGCAATT 1259
DB 1158 ATTTTGAAGCGGCAAAACCTGTCCCTGGAAACAGATTCTCTCTTTATGGCATT 1210
```

RESULT 9

```
US-10-424-599-93080
; Sequence 93080, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 93080
; LENGTH: 2293
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_55065C.1
US-10-424-599-93080
```

```
Query Match 35.4%; Score 614.6; DB 7; Length 2293;
Best Local Similarity 71.8%; Pred. No. 7.1e-171;
Matches 821; Conservative 0; Mismatches 314; Indels 8; Gaps 1;

QY 333 TCTTCAAGAACAAATGAAGCGTCAGGACCAAGAGTCCGACAAATTTAAGGGCTCTTTGTCA 392
DB 620 TCTTTCAGAGAGCAGAAACGTCGAGACCAAGAAATGTGGACAAACAAAGTCTTTGTACA 679
QY 393 GGATCTTTGAAGTAAAGGCGCATAAAAAGTTGATTCGGAATGTACAGATGCCAGTGGCTGC 452
DB 680 AGATCTTGAAGAAAAAGACCTTCAGAGCGCTGATTGATAAAGTGCAGGTTCCGGTGGCTGC 739
QY 453 TGTAGTTGTTATGGCTTGCATTCGGCTGATTTACCTGGAAGAAAGCTATTAAATCCATCTT 512
DB 740 TGTGTGATCATGGCATGTAATCGTCTGCTAATACCTGGAGAGGACTATTAAATTTCTGTCAT 799
```

```
Qy 513 AABATACCAATATCTGTTGGTCAAAATATCCCTCTTTTCATATCCAGGATGGATCACA 572
Db |||
Qy 800 AACTTCAGCGGTGAATCATCTTGAAGATGACATTTGAATAGCAACGGATTTCTTCAAT 859
Db |||
Qy 573 TCCTGATGTCAAGAGCTTGTCTTTCAGCTATGATCAGCTGACGTATATGACGACCTTGGGA 632
Db |||
Qy 860 TATTGTAAGCTGCAGCGAGTCTTCTTGAGATGATAAATCCATTTATGGCTGTTTCTT 919
Db |||
Qy 633 TTTTGAACC-----TGTGCATCTGAAGAAGACAGGGAGCTGATGCAATCAAAA 684
Db |||
Qy 920 ATGGAATGCCACTTGGATTTTGAACCTGAACGGCTGGAGAGTTAACTGCTTACTACAAA 979
Db |||
Qy 685 ATTGCAGCTATTACAAGTGGCATTCGATCAGCTGTTTACAGCATATAATTTAGCGGT 744
Db |||
Qy 980 ATTGCAGCTATTACAAGTGGCTCTAGATCAACTGTTTACAAAGCATAACTTCAGCGGT 1039
Db |||
Qy 745 GTTATCATCTAGAAGATGATGAAATTCGCCCTGATTTTTCGCTTTTTCAGGGCT 804
Db |||
Qy 1040 GTATCATCTTGAAGATGATGAAATAGCACTGATTTCTTGGATTTTGAAGCT 1099
Db |||
Qy 805 GGAGCTACTCTTTCAGACAGACAAGTCGATTATGGCTATTTCTTTCGGAATGACAAT 864
Db |||
Qy 1100 GCAGCGACTCTCTTGAAGGATAAATCCATTTATGGCTGTTCTCATGGAATGACAAT 1159
Db |||
Qy 865 GGACAAATGCAGTTTGTCCAAGATCCTTATGCTCTTTTACCGCTCAGATTTTTCGCGGT 924
Db |||
Qy 1160 GGACAAAGCAGTTTGTACATGATCCATATGAATCTTTATCGCTCTGACTCTTTCCTGGA 1219
Db |||
Qy 925 CTGAGTGGATGCTTTTCAAAATCTACTTGGGACAAATTTATCTCCAAATGGGCCAAAGGCT 984
Db |||
Qy 1220 TTAGATGGATGCTGGCCAGATCTACATGGATGAGCTATCACCAATGGCCAAAGCT 1279
Db |||
Qy 985 TACTGGGACGACTGGCTAAGACTCAAGAGAAATCACAGAGTTCAGCAATTTATTCGCCCA 1044
Db |||
Qy 1280 TACTGGGATGACTGGTTGAGACTTAAAGAGAAATCACAAAGGACGACAGTTTATCCGGCC 1339
Db |||
Qy 1045 GAAGTTTCGAAACATATAATTTTGGTGAGCATGTTCTAGTTTGGGCGAGTTTTCAG 1104
Db |||
Qy 1340 GAAGTATGCAGAACATATAATTTTGGTGAGCATGGTCTAGTTTGGGACAGTTTTCAG 1399
Db |||
Qy 1105 CAGTATCTTGAGCCAAATTAATGATGTCCAGGTTGATTGGAAATCAATGGACCTT 1164
Db |||
Qy 1400 CAATTTCTTGAGCCAAATCAAGCTGATGATGTCAAGTTGATTGGAAATTAATGGATCTG 1459
Db |||
Qy 1165 AGTTACCTTTTGGAGGACAAATTAAGTGAACACTTTTGGTGACTGGTTTAAAGGCTAAG 1224
Db |||
Qy 1460 AGCTATTTACTGGAGGATAAATATCTATGCACCTTTCGGAACGTTGTTAAGAAAGCTACA 1519
Db |||
Qy 1225 CCCATCCATGAGCTGATGCTCTTGAAGCAATTTAATAGATGATGATGTCGCTATT 1284
Db |||
Qy 1520 CCTGTCTATGAGCTGACATGGTCTTAAGGCATCTAATATAGATGGGATGTGCGCATC 1579
Db |||
Qy 1285 CAGTACAGAGATCAACTAGCTTTGAAATATCGCAGGCAATTTGGCATTTTGAAGAA 1344
Db |||
Qy 1580 AAATATAAGATCAGTCACTTTGAAACAACTGCTCACCAATTTGGTATATTTCAAGAG 1639
Db |||
Qy 1345 TGAAGGATGTGTACCAAGTGCAGCATATAAAGGAATPAGTATTTCCGGTACCAAAACG 1404
Db |||
Qy 1640 TGAAGGATGTGTGCCCGGACAGCATATAAAGGAGTAGTTCGTTTTCAGATATCAAAAC 1699
Db |||
Qy 1405 TCCAGAGCTGATTTCCCTTGTGGCCATGATTTGCTTCAACAACTCGGAATTTGAAGTACT 1464
Db |||
Qy 1700 TCAGACGATATATCTTGTGGTCCAGAACTTTTGAAGCTACTTCAGATCGAAGAGTCT 1759
Db |||
Qy 1465 TAA 1467
Db 1760 TAA 1762
```

RESULT 10

US-10-739-930-5511/c

; Sequence 5511, Application US/10739930

; Publication No. US20040216190A1

```
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 5511
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1898)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-CLUSTER98_776
US-10-739-930-5511
```

Query Match 33.5%; Score 581.2; DB 8; Length 1898;
Best Local Similarity 66.4%; Pred. No. 5.3e-161;
Matches 887; Conservative 0; Mismatches 429; Indels 19; Gaps 3;

```
Qy 147 CTGTGATTTCCGGTACCTCTCATCTTGGCTGCTGCGCTTCATCTACACAGATGCG 206
Db |||
Qy 1778 CTGCGACATCCGGCTCTCTCTGTCGCCGCCGCTGCTTTCATCTACATCCAAGTGG 1719
Db |||
Qy 207 GCTTTTGGGACACAGCTCAGAAATATGCAGATCGCTTGTCTGCTGCAATTTGAAGC 260
Db |||
Qy 1718 GCTGTTCTCCACCACCTCTTCTTCTTCTTAATCCCACTTTCAGACCCGAGC 1659
Db |||
Qy 261 -----AGAAATCATTTGTACAAAGCAGACAGATTTGCTTATTTGACAGATTTAGCCT 311
Db |||
Qy 1658 CTTGCGGGAAGCAGAGAAATGCGAGAGCCAGCTGCACGCCATGANGCAGCCAGGTCAGCG 1599
Db |||
Qy 312 GCAGCAAGGAGAAATAGTTGCTCTTCAAGAACAAATGAAGCGTCAGAGCCAGGAGTGGCG 371
Db |||
Qy 1598 CCAGCAGGAGAAATCGCCGCACTTGAAGAGATGAAGGTACGCCAAGAGAGGCGGTGT 1539
Db |||
Qy 372 ACAATTAAGGGCTCTTGTTCAGGATCTTGAAGTAAGGGCATAAAAAGTTTGTATCGGAA 431
Db |||
Qy 1538 GCAGCTCAAGATTTTAATCCAGATCTTGAAGAGAGAGTTTACAGACCTTAACAAACA 1479
Db |||
Qy 432 TGTAAGATGCCAGTGGCTGCTGATGTTGTTATGCTTGCATTCGCAATCGGGCTGATTTACCTGGA 491
Db |||
Qy 1478 GAAGTGGTTCCTGTCGCGCTGTTGTCATATGCTTGCATATGCTTGCATATGCA 1419
Db |||
Qy 492 AAAGACTATTAATCCATCTTAAATACCAATATCTGTTGCTGCAAAATATCTCTTTT 551
Db |||
Qy 1418 GAGGACAGTGGAAATCTATCTCAAGTACCAAGAGCAGTTGCTTCAAGAGTTTCCACTATT 1359
Db |||
Qy 552 CATATCCAGGATGGATCAGATCCTCATGTTCAGAAAGCTTGTCTTTCAGCTATGATCAGCT 611
Db |||
Qy 1358 TATATCACAGGATGGAACAAATGGAAGATTAAGAAAGAGCTTTGAGTTACAAAT 1299
Db |||
Qy 612 GAGCTATATGACACA-CTTGGATTTTGAACCTGTGCTGCTGCAATCTGAAAGCAGGAGGAGCTGA 670
Db |||
Qy 1298 AACATTTATGACAGATGTGGATCTTTGAGCTCTTGGCCTGTGGCCTGAAAGACCAAGAGAAACG 1239
Db |||
Qy 671 TTGCATCTACAAATTTGACGCTCATACAGTGGGCATTTGGATGAGCTGTTTACAGC 730
Db |||
Qy 1238 TTGCATATTACAAAGATAGCTAACCACTATAAATTTGGGCTTTGGATGAGCTATTATTAAAGC 1179
Db |||
Qy 731 ATAAATTTAGCCGTGTTTATCATACTAGAAGATGATATGGAATTTGCCCTGATTTTTCG 790
Db |||
Qy 1178 ATGATTTTTCGTCAGTAAATCAATCTCGAAGATGACATGAGATCGCCCGAGATTTCTTTCG 1119
Db |||
Qy 791 ACTTTTTCGAGGCTGAGCTACTCTTTCTTGACAGAGACAAGTCGATATAGGCTATTTCCTT 850
Db |||
Qy 1118 ACTACTTTGAGGCTGCAGCGGAAATTTACTTGACACTGACAAGACAATAATAGCTGTTTCAT 1059
Db |||
Qy 851 CTTGGAATGACATGACAAATGCAGTTTGTCCCAAGATCCTTATGCTCTTTTACCGCTCAG 910
Db |||
```



```

1058 CTTGGAATGATTAAGGCAAGAGTCTGTTTATGACCCAAAGGCTCTTTACCGTTCGG 999
911 ATTTTTCCTCCGGTCTTGGATGATGCTTTTCAAAATCTACTTGGGACGAATATCTCCAA 970
998 ACTTCTTTCCTCCGGGCTTGGATGATGCTTAAACGAAGTCAACATGATGAGCTGTGCGCAA 939
971 AGTGCCAAAGGCTTACTGGGACGACTGGCTAAGACTCAAGAGAAATCACAGAGGTGAC 1030
938 AGTGCCCAAGGCTTATTTGGGATGACTGGGTGAGGCTTAAGAGGATCACAGAGATCGG 879
1031 AATTATTTTCGCCCAAGGCTTTCGAGAACATATAATTTTGGTGAGCATGGTTCTAGTTGG 1090
878 AGTTTATTTTCGCCCAAGGCTTTCGAGAACATATAATTTTGGTGAGCATGGTTCTAGTTGG 819
1091 GGCAGTTTTTCAAGCAGTATCTTGAGCCAAATTAACATAATGATGTCAGGTTGATGGA 1150
818 GACAATCTTTGATCAATACCTTGAACCTATCAAGTTAAATGATGCTCATATTGACTGGA 759
1151 AGTCAATGAGCTTGTATACCTTTTGGAGGACAAATTAAGTGAACACATTTTGTGACTTGG 1210
758 ATTCGAGGACTGAGCTACCTCAAGGAGGACAAATTTTGAACCAATTCGGGAAAGACG 699
1211 TTAAGAGGCTAAGCCCATCCATGAGCTGATGCTCTTGAAGGCAATTTAAACATAGATG 1270
698 TGGCTAGCGCCACACCTGTGATGATCGATGCTCTTGAAGCCCAATCTGGATG 639
1271 GTGATGTCGATATTCAGTACAGAGATCAACTAGACTTTGAAATAATCCAGCGCAATTTG 1330
638 TGGAGCTAAGGATTCAGTATGACATCAGGCGCACTTTGAGCGTATAGCTCGGCAATTG 579
1331 GCATTTTGAAGATGGAAGGATGCTTACACGTCGAGCATATTAAGGAATAGTAGTTT 1390
578 GAATATTTGAAGATGGAAGGACGCTGTTCCACGGGCGGCTTACAAAGCGTGGTGGTGT 519
1391 TCCGGTACCA---AAGCTCCAGACGCTGATTTCTTTGTTGGCCATGATTCGCTTCAACAC 1447
518 TCCGGTACAAAGGCGAGCGGAGACGATATACCTGGTGGTCCGACTCTTCCGCAAC 459
1448 TCGGAATGAAGATA 1462
458 TCGGGGTTTAGCTTA 444

```

```

RESULT 11
US-10-424-599-93082
; Sequence 93082, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 93082
; LENGTH: 1414
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(1414)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_55067C.1
US-10-424-599-93082

```

Query Match 33.2%; Score 577; DB 7; Length 1414;
Best Local Similarity 72.3%; Pred. No. 7.8e-160;

```

Matches 814; Conservative 0; Mismatches 261; Indels 51; Gaps 3;
Qy 139 AAGTTTTCCTGCTGATTTTCCTGCTACCTCTCATCTTGGCTGCTGCTGCTCATACACA 198
Db 333 AAGTTTTCCTGCTGATTTTCCTGCTACCTCTCATCTTGGCTGCTGCTGCTCATACACA 392
Qy 199 CAGATGCGGCTTTTGGGACACAGTCAAGATATGACAGATCGCTTGTGCTGCAATGAA 258
Db 393 CAGATGAGGCTTTTGGCAACACAAATCAGAAATATGCTGATCGCTCGCTGAGCTATCGAA 452
Qy 259 GCAGAAATCATTTGTAACAGCCAGACAGATGCTTATTCACAGATTTAGGCTTCAGCAA 318
Db 453 GCTGAAACCATTTGTAACAGCTCAACCGGATCGCTGATTTGATCAGATTTAGCTTCAACAA 512
Qy 319 GGAAGAATAGTTGCTCTTTGAAGA-----ACAAATGAAGCGCTCAGGACAGGAGTGC 371
Db 513 GGAGCAATTTGGCCCTTGGAGAGAAACGTTAGAGCAGAAACGTCGAGACCAAGAAATGTGG 572
Qy 372 ACAATTAAGGCTCTTGTTCAGGATCTTGAAGATGAGGACATATAAAGTTGATCGGAA 431
Db 573 ACAACAAAGTCTCTTGTACAAGATCTTGAAGAAAGAGACCTGACAGAGGCTGATGTGATA 632
Qy 432 TGTAACAGATGCGAGTGTCTGTAGTTGTATGGCTTGAATCGGCTGATTAACCTTGA 491
Db 633 AGTGCAGGTTCCGGTGTCTGTGTGATCATGGCATGTAATCGTCTGATTAACCTTGA 692
Qy 492 AAAGACTATTAAATCCATCTTAAATACCAATATCTTTCGCTCAAAATATCTCTTTT 551
Db 693 GAGGACTATTAAATCTGTATTAATTAACAAAGGCCCATTTCTTCAAGATATCTTTATT 752
Qy 552 CATATCCAGGATGATCATCTGTGATGTCAGAAAGCTTGTGCTTGTGAGCTATGATCAGCT 611
Db 753 TGTATCTCAGGATGATCAAACTCGAAACTGGAGATAAACTGAGGGAC-----TAT 801
Qy 612 GACGTATATGCGACACTTGGATTTTGAACCTGTGCATCTGAAAGACACCGGAGGAGTGTAT 671
Db 802 -----AAATTTGCAATTTTGCCTAAAGAA-----TAT 828
Qy 672 TGCATCTACAAATTTGCACGTCATTAACAAGTGGGCTTGGATCAGCTGTTTAAACAGA 731
Db 829 TCAATTTGATTCGGTTATCAGGTCAATTAACAAGTGGGCTTGTAGTCAACTGTTCTACAAGCA 888
Qy 732 TAATTTTAGCGCTGTTATCATCTAGAAATGATATGAAATTTGCCCTGATTTTGTGA 791
Db 889 TAACCTTACGCGCTGTTATCATCTTGAAGATGATGAAATAGCACCTGATTTCTTTGA 948
Qy 792 CTTTTCCTGAGCTGAGTACTCTTTCAGAGACAGAGTCAAGTATGCTGATTTTCTTC 851
Db 949 TTAATTTGAAGCTTGCAGCGAGTCTCTTTGAGAAGGATAAATCCATTTATGGCTGTTTCTTC 1008
Qy 852 TTGGAATGACAAATGACAAATGACAAATGCTTATGCTTATGCTCTTTACCGCTCAGA 911
Db 1009 ATGGAATGACAAATGACAAAGAGGTTTGTATGATGATCAATGATTTATCGCTCTGA 1068
Qy 912 TTTTTCCTCCGGCTTTCGATGAGTCTTTCAAAATCTACTTTGGGACGAATTTATCTCCAAA 971
Db 1069 CTTCTTCCTGATTTAGGATGAGTGTCTGCGCAGATCTATCATGGAATGAGCTATCAACAAA 1128
Qy 972 GTGGCCAAAGGCTTACTGGGACGACTGGCTAAGACTCAAGAGATCAAGAGGTCGACA 1031
Db 1129 ATGGCTAAAGCTTACTGGGATGACTGGTGTGAGACTTAAAGAGAAATCAAAAGGACGACA 1188
Qy 1032 ATTTATTCGCCAGAGTTTCAGAAACATATAATTTTGTGAGCATGTTTCTAGTTTGGG 1091
Db 1189 GTTTATTCGCCGCGGAGGATGACAGAAACATATAATTTTGTGAGCATGTTTCTAGTTTGGG 1248
Qy 1092 GCAGTTTTTCAAGCAGTATCTTTGAGCCAAATTAACATAATGATGCTCAGGTTGATGGA 1151
Db 1249 ACAGTTTTTAAAGCAATCTTTGAGCCCAATCAAGCTGAATGATGATCAAGTTGATGGA 1308
Qy 1152 GTCAATGAGCCTTACTGCTTTTGGAGGACAAATTAAGTGAACACTTTGGTGGTACTTGT 1211
Db 1309 ATTAATGGATCTGAGCTATTTACTGGAGGATAAATTAATCTATGCACTTNTGCAATGTTAT 1368

```


778 TATCTGAGTATCAGCATCAGTGGCTTCAAGTTCCATATTTATCTCAGATGG 837
567 ATCATCTCTGATGTCAGAAAGCTTGTCTGAGCTATGATCAGCTGAGTATATGACGA 626
838 AATAATGGAGAGTGAAGAAAGCTTGTGATTTATAATGAATGAATATATGACAT 897
627 CTGAGATTTGAACTGTGCTATCTGAAGACCGAGGAGCTGATGTCATCTACAAAT 686
898 CCAATGAGTATATCAGGAAACTCTAGGTCCAATGATATATACTGAGCAGTGAGACGA 957
687 TGCAGCTCAT--TACAAGTGGGCAATGGATCAGCTGTTTTCACAGCATAAATTTAGCCGT 744
958 AGGCCATATCATACCCAACTTGTGATGATTTGTTGAAAGGAAAGCTTGTAGCCAC 1017
745 GTTATCATCTAGAGATGATGAAATGTGCCCTGATTTTGTGACTTTTTCAGG-- 802
1018 TGTACTGTACTTCCCTCCACATAACATAGGCTGAGCAGTCCGATATTTGTAAGA 1077
803 ---CTGGAGCTACTCTTTCAGACAGACAACTCGATATGCTATTTCTTCTGGAATG 859
1078 ACATGGGAATATTTTTCAGCTAAGACTCTCAGTGAACATTAACAATGGGCCCTTGGATGA 1137
860 ACAATGGA--CAATGCGAGTTCTCCAAAGATCCTTATGCTCTTTACCGCTCAGATTTTTT 918
1138 GCTATTCATCAACACACTTTTGGCGAGTATATAATCTGGAAGATGATATGAGATCGC 1197
919 CCGGCTTTGAGATGATTTTCAAAATCTACTTGGGAGCAATTTATCTCAAAAGTGGCA 978
1198 CCTGATTTTGTACTACTTGGAGGCTGAGCTAAATTTACTTTGATAACGATAAGACAAT 1257
979 AAG----- 982
1258 CATTGCGCGTTCTTCTGGAATGACAAATGGAACAAAGCAATTTGTTAGACCCAACTT 1317
983 ---CTTACTGGGAGCTGCTGAAGACTCAAGAGAAATCAAGAGGTGCAAAATTTATTGCG 1040
1318 TACTTATTTGGATGATCTGGTGAGCTAAAGAGGTACACAGACCGGCAATTTATTGCG 1377
1041 CCCAGAGTTTGCAGAACATATAATTTTGGTGAAGATGTTCTAGTTTGGGCGAGTTTTT 1100
1378 CCGGAGGCTGCAGAACATACAAATTTTGGCGAACTGCGTCAAGCATGGGGCAATTTCT 1437
1101 CAAGGATTTTGGACCAATTAACATAATGATGCTCAGGTTGATGGAAGTCAATGGA 1160
1438 CAGACAGTATTTGGAACCAATCAAGCTAAATGATGCCCATATCAAGTGGAATTTCTGAGGA 1497
1161 CCTTAGTTACCTTTTGGAGGACAAATACGTGAACACTTTGTTGACTTGTGTTAAAGGC 1220
1498 CTTGAGCTACCTCAAGAGGACAAAGTTCTGATCCAAATTTGGGAAGACGTCCTAGTGC 1557
1221 TAAGCCCATCCATGGAGCTGATGCTCTTGTGAAGCAATTTAACAATAGTGGTATGTCG 1280
1558 CACCCCTCTCCATGGATCTGATGCGCGTTTGAAGCCCAATATGATGGATGGGATGAAG 1617
1281 TATTCAGTACAGAGATCACTAGACTTTGAAATATCGACGCAATTTGGCAATTTTGA 1340
1618 GATCCAGTATACGATCAGGAAGACTTTGAGCGGATAGCTCGTCAATTTGGAATATTTGA 1677
1341 AGAATGGAAGGATGTTTACCACTGACGATATAAAGGAATAGTATGTTTCCCGGTACCA 1400
1678 AGAATGGAAGGATGGCAATTTCCAAGACAGCTTATAAGGATAGTATGTTCCCGGTACCA 1737
1401 ---AAGCTCAGACGCTGATTTCTTGTGGCCATGATTCGCTTCAACAACCTCGGAAT 1455
1738 GAGTAGTCGAGACGATATACCTCGTTGTCGGATTTCCCTCAGTCAGCTTAGGTTT 1795

RESULT 14
US-10-767-701-5996
; Sequence 5996, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:

APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
FILE REFERENCE: 38-21(5353)B
CURRENT APPLICATION NUMBER: US/10767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 5996
LENGTH: 769
TYPE: DNA
ORGANISM: Sorghum bicolor
FEATURE:
NAME/KEY: unsure
LOCATION: (1)-(769)
OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS10971_1
US-10-767-701-5996
Query Match 12.8%; Score 222; DB 7; Length 769;
Best Local Similarity 64.7%; Pred. No. 1.6e-54;
Matches 330; Conservative 0; Mismatches 180; Indels 0; Gaps 0;
QY 147 CTGTGATTTTCGGTACCTCTCATCTTGGCTGCTGTGCGCTTCATCTCATACACAGATCGG 206
DB 259 CTGGACCTCGCT 318
QY 207 GCTTTTTCGACACAGTACAGATATGACAGTGCCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 266
DB 319 GGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 378
QY 267 TCATTGTACAAGCCAGACAGATTTCTTATTTGACAGATTTAGCTTCGACGAGGAAAGAT 326
DB 379 TCAATGACACCTGCTCACTAAAGTCTCTGATCGATCGATCGATCGATCGATCGATCGATCGAT 438
QY 327 AGTTGCTCTTGAAGAACAAATGAAGCGTCAGGACGAGGTCGCGACGAGTCCGACAAATTAAGGCTCT 386
DB 439 TGTAGCATTTGGAAGAGATTAAGGTACGCAAGATGAAGAACGCGCACATCTGAAGATTT 498
QY 387 TGTTCAGATCTTGAAGTAAAGGATTAAGAAAGTTGATCGAAATGTCAGATGCCAGT 446
DB 499 GATAAAGGATCTTGAAGAAAGGAGTGTGCAAGAGTCTAGACAAATTAATGTTGCTCTGT 558
QY 447 GCT 506
DB 559 TGTCTGCT 618
QY 507 CATCTTAAATAACCAATATCTGTTGCGTCAAAATATCTCTTTTTCATATCCAGGATGG 566
DB 619 TATCTGAAATATCAGACACAGTTGCTTCAAGTTTTCGCTTTTATATCTCAGATGG 678
QY 567 ATCATCTCTGATGTCAGGAAAGCTTGTGAGCTATGATGATGATGATGATGATGATGATGATGATGAT 626
DB 679 AGCAATGGAGCTGTCAAAAGGAAAGCTTTGGAATATAAACAATAAATAATATATGACGA 738
QY 627 CTTGATTTTGAACCTGTCATCTGAAAG 656
DB 739 TGTGATCTTGAACCTGTCGAAAGTGAAG 768
RESULT 15
US-09-923-876-5569
; Sequence 5569, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Kamigaki, Laura Y. (Ito)
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
; FILE REFERENCE: PL-0012-1 CON
; CURRENT APPLICATION NUMBER: US/09/923,876

```

; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 5569
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700457259H1
US-09-923-876-5569

```

```

Query Match      10.5%; Score 182; DB 3; Length 278;
Best Local Similarity 78.4%; Pred. No. 6.5e-43;
Matches 218; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 625 CACTTGGATTGTAACCTGTGCATACGAAAGACGAGGGAGCTGATTGCATACATAAAA 684
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1 CATGTGGATCTTGAACCTGTGAAACTGAAAGGCCAGGAGAAATTGACAGCATATTACAAG 60

Qy 685 ATTGCACGTCAATACAAAGTGGGCTTGCATCAGCTGTTTACAAAGCATATTTTAGCCGT 744
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 ATTGCTAAACATTATAAGTGGGCTTGGACAACCTATTTCATTAAACATAATTTTGCTCGA 120

Qy 745 GTTATCATCTAGAAGATGATGGAATTCGCCCTGATTTTGTGACTTTTGTGAGGCT 804
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 GTATCATCTAGAAGATGATGGAATTCGCCCTGATTTTGTGACTTTTGTGAGGCT 180

Qy 805 GGAGCTACTCTTCTTGACAGAGACAAGTCGATTATGCGCTATTCTTCTTGGAAATGACAAT 864
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
181 GCAGCTAAACTACTTGAATGACAAGACGATTATGGCTGTTTCTGCTTGGAAATGACAAT 240

Qy 865 GGACAAATGCAGTTTGTCCAAGATCCTTATGCTCTTTA 902
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 GGGCAGAGCAGTTTGTTCACGACCCAAAGCTCTTTA 278

```

Search completed: December 14, 2005, 13:30:43
Job time : 1313 secs

November 2005

Published_Applications_Nucleic_Acid and Published_Applications_Amino_Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).

Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

page blank (uspio)